LEVEL III

V1 - A056400
V3 - A056469

①

⑥

Army Science Conference Proceedings,

Held on

20-22 June 1978.

Volume II.

Principal Authors G-M

⑪

1978

⑫ 429p.
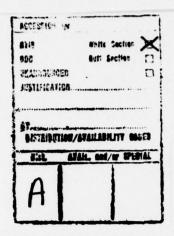
This document has been approved for public release and sale; its distribution is unlimited.

78 06 09 080

408 992

GAMBINO

RE: Classified References--
    Document should remain for unlimited
distribution. Per: (Mrs.) Anne G. Taylor,
ARO.

AN EXPERIMENTAL DIGITAL INTERACTIVE FACILITY (U)

*LAWRENCE A. GAMBINO, MR.
US ARMY ENGINEER TOPOGRAPHIC LABORATORIES
FT. BELVOIR, VA   22060

INTRODUCTION

The U.S. Army Engineer Topographic Laboratories (USAETL),
Fort Belvoir, Virginia, has developed an extensive, interactive
digital image processing facility to efficiently conduct research in
areas such as military intelligence, digital product generation for
weapon systems, and digital mapping. The facility is called the
Digital Image Analysis Laboratory (DIAL), and it began to evolve from
a general purpose computer-oriented system approximately 4 years
ago. It is supported by the Defense Mapping Agency (DMA), the Corp
of Engineers, and DARCOM's Army Space Program Office (ASPO).

The research effort in digital image processing began in a
modern computer environment 6 years ago. As experience was gained,
it became apparent that it was impractical to produce digital pro-
ducts in this environment because of the computer's inefficiency re-
lative to existing production equipment using electro-optical tech-
nology. A classic example of this in the mapping sciences is seen in
the development of stereocompilation equipment during the past 20
years where the computer is never used to perform the basic correla-
tion function. Therefore, it was necessary to create a digital sys-
tem which would at least indicate its competitive qualities relative
to current production systems, and it placed the research on a higher
plane as opposed to simply testing algorithms on general purpose
computers.

Data volume is a characteristic of digital image processing
and it significantly impacts the resources of conventional computer

systems by locking out other tasks in a multiprocessing computer environment. We wish to use this characteristic to distinguish between digital image processing and computer graphics where, generally speaking, the former implies at least six times as much data is being manipulated as compared to applications involving computer graphics. At least 64 grey levels are represented on a good quality photograph which requires six bits to digitally represent a spot (pixel) while graphics generally require only one bit to represent lines.

These adverse consequences upon conventional computer systems drive digital image processing research in three directions. One lies in the area of efficient algorithm development, such as the Fast Fourier Transform (FFT). Another is in the direction of efficient system software development where lack of concern has caused execution times of up to ten times longer for the same algorithm on systems with similar hardware. Finally, digital image processing spurs research in digital architecture which includes display technology. As we describe DIAL and various applications in this paper, we see that all three areas have been considered in the system development.

HARDWARE

We will not elaborate in detail about the hardware since much of it involves standard computer components. However, we will point out those nonstandard hardware items which were acquired to build the interactive system. This comment implies that the software and hardware together make up a system which is not commercially available through the standard computer manufacturer and that there are no similarities between this software system and commercially available interactive graphics software packages.

Figure 1 is a block diagram of the DIAL system. It consists of three subsystems: the host control processor (CDC-6400), the associative array processor (STARAN), and the image softcopy/hardcopy/digitizing subsystem (COMTAL 8300, DICOMED D-56, D-47, D-36; PDP-11/50) [1]. Each subsystem operates either as a stand-alone or together with other subsystems as an entire system. Digital interactive image processing goes on concurrently with testing and software development on the host processor. Imagery can be passed from one subsystem to another via commands from the Tektronix terminals whereby processing might take place in STARAN and the results viewed on the softcopy system in a matter of seconds.
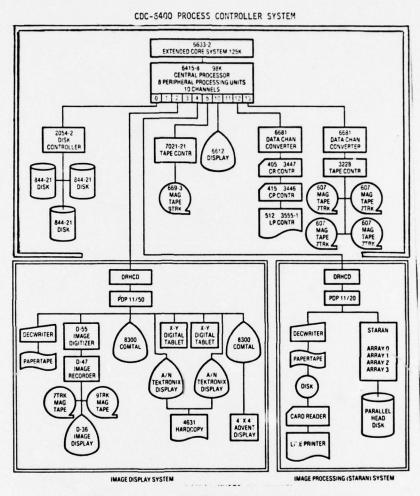
CDC-8400 PROCESS CONTROLLER SYSTEM

Figure 1. DIAL'S THREE SUBSYSTEMS

STARAN. Developed and built by Goodyear Aerospace Corporation, Akron, Ohio, the STARAN represents one of the aforementioned nonstandard hardware components of the interactive system. It is the parallel processing part of the system and only those image processing functions which can be executed faster using its digital, parallel architecture, are transmitted to it for processing. All applications implemented on STARAN have required a considerable rethinking of the problem in order to take full advantage of its architecture, which is unique among all digital processing systems.

GAMBINO

We will discuss only the characteristics of the arrays of STARAN in this paper. Additional information can be found in [2]. There are four arrays in the STARAN interfaced to the interactive system. Each of the four arrays is composed of 256 words, each having 256 bits for a total of 1,024 words and 262,144 bits. Each array contains 256 simple processing elements. Therefore, 1,024 processing elements are available to perform an operation simultaneously. Further, each processing element acts on an independent data stream, thereby offering the theoretical possibility of acting simultaneously on 1,024 independent data streams. Thus far, Input/Output restrictions have kept us from taking full advantage of this very large processing bandwidth. A simplified diagram of one array is shown in Figure 2. The arithmetic operations are bit serial but
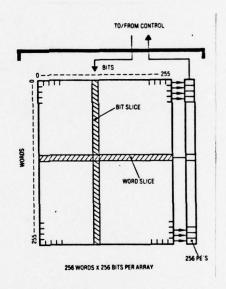

Figure 2. Associative array

word parallel -- architecturally the inverse of conventional computers.

Another unique feature of the array is its content-addressed (associative) memory. This allows identity of all elements meeting a certain criteria in a single memory access whereas the conventional computer searches its data file one element at a time. This feature will find widespread use in data base management and lead to a logical structure of data not now utilized in commercially available data

management software packages. As a matter-of-fact, we are investigating a relational data structure as opposed to the network and hierarchical structures prevalent in current data management systems. The relational structure is suited to associative memories and it also more nearly represents the way people usually structure data; namely, as flat files like in telephone directories.

CHANNEL COUPLERS. These hardware devices represent other nonstandard computer components and were procured to interface the host processor, which is the Control Data Corporation (CDC)-6400 computer, to the STARAN and image display subsystem. The channel coupler permits up to four digital image formats to be programmed and these formats are transferred across these channels at approximately 4.8 megabits/second. However, because of system software overhead, the effective rates are approximately 2.5 megabits/second. These channel couplers were built by the Digital Equipment Corporation (DEC) and detailed information concerning them is available elsewhere [3].

PDP-11/50 COMPUTER. The PDP-11/50 provides the interface between the image display subsystem and the CDC-6400 computer through the channel coupler. In addition to serving as this interface, it is used as the real-time controller for all devices connected to it. It also allows local processing of imagery as opposed to performing operations on the CDC-6400 computer or STARAN. Operations such as grey-scale remapping and pseudocolor encoding are performed "locally" since these kinds of image processing operations do not require the power of the other computers.

Many of these simpler operations can occur concurrently with processing on the CDC-6400 or STARAN. For example, while the PDP-11/50 is remapping grey shades, the CDC-6400 can be pseudocolor-encoding grey shade data stored in the display system.

DIGITAL IMAGE DISPLAYS. The image display units consist of two COMTAL 8300-SE systems featuring refresh storage of three 512 x 512 1-bit graphic overlays in each system. We will see the graphics capability in the photos reproduced later in this paper. The reader should keep in mind that these displays are high resolution and are much more expensive than the home television variety. High resolution is a necessary factor if serious photo interpretation for military intelligence operations is to take place in a digital environment as compared to work at the conventional light table.

In concluding this section on hardware, we must state that the

characteristics of each of the subsystems have been presented in only very general terms and only very briefly. We have not said much about the CDC-6400 computer itself, but let it suffice to say that it acts as the controller for the interactive system and that it has some features which are well suited to digital image processing. All told, five different computers make up the interactive system. Although each of these computers is not explicitly mentioned in this paper, they make up parts of the larger computers. For example, the CDC-6400 has eight peripheral processing units (PPUs) which are 12-bit, 4,096-word minicomputers. These control all the peripheral hardware and much of the operating system software. Two of these PPUs are dedicated to nonstandard use in that one is used for interfacing STARAN to the system and the other for interfacing the PDP-11/50. These two units are used to accomplish all the control and data transfers between the subsystems. The other computer not previously mentioned is a PDP-11/20 used to control the internal operation of STARAN. Therefore, the PDP-11/50, PDP-11/20, PPU, CDC-6400 and STARAN make up the five different computers used in the development of the digital interactive system. Needless to say, a very sophisticated system software development program was undertaken to make the interactive system operational.

SOFTWARE

Two bodies of software were developed for use by non-system personnel. It is via these software systems that users and applications personnel develop their programs to perform intelligence and mapping functions. Relieving these people from system software concerns should speed research in digital image processing. The two bodies of software are called the Menu system and the FORTRAN interface system.

MENU SOFTWARE. This is referred to as control software since it provides the user a capability for exercising the computer system, processing functions, and hardware peripherals. It is designed to be a tool since it is highly user-oriented and, as such, it has been tailored to the needs of the user who has customarily utilized light tables and microscopes for the extraction of information from photography. With MENU software, the non-computer-oriented is led through such image processing operations as search, scale, rotate, translate, generate subimage, reformat, remap grey levels, etc. Some of the display operations include display entire image or subimage, erase screen, save image, edit images, etc. Generally speaking, many application programs exist within the MENU system. There are approximately 80 different operations in this system and it

GAMBINO

is designed to assist a photo interpreter in making decisions
quickly and accurately.

FORTRAN INTERFACE SOFTWARE. This body of software allows the
user to access his own FORTRAN programs using system routines for
input/output and to execute them from the Tektronix terminal. It
also allows communication with the STARAN via FORTRAN -- callable,
routines in addition to allowing use of the Tektronix and COMTAL dis-
plays via FORTRAN-like statements. Not much else will be said of
this software other than to emphasize the fact that implementation
of digital image processing applications can be accomplished quickly
since the system software handles all the data management thereby
relieving the researcher of this burden. He is free to quickly de-
velop applications in this FORTRAN environment and speed up the re-
search in digital image processing.

APPLICATIONS

We will present results of several functions which a photo
interpreter performs in carrying out military intelligence opera-
tions along with a few simple mapping functions. These results were
obtained by writing application software using the FORTRAN interface
software system.

All of the following photographs were taken on the face of the
COMTAL displays using a polaroid camera and do not reflect the true
resolution of these displays.

Military Intelligence-Scroll/Magnify/Mensurate/Target. These func-
tions are depicted in several of the following photographs. However,
they were performed on the same image as the image was being scrolled
on the displays by the photo interpreter. It indicates the multi-
function capability of the system under operator control.

Scrolling is an important data handling function since it
allows a photo interpreter to quickly scan a large image at reduced
resolution and pick out areas for closer investigation at full
resolution. The technique is analogous to moving map displays ex-
cept here we are moving a digital photograph and much more data. For
example, Figure 3 is a reduced resolution photo of a 2,048 x 2,048
pixel, 8 bits per pixel image stored in the interactive system data
base. It contains approximately 32 million bits of information and
not all of this data can be displayed on the COMTAL screens because
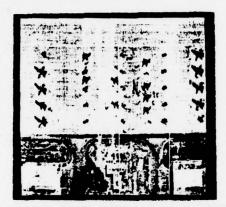they are limited to 512 x 512 pixels, 8 bits per pixel. Therefore,

Figure 3.   Reduced Resolution/Scroll

the scrolling operation requires that not more than every fourth
pixel be fetched from the data base, a line at a time, and moved to
the COMTAL displays. The software required for this seemingly
simple operation is complicated if it is to be made efficient and a
simple, brute-force approach only slows the process and defeats the
objectives of an efficient interactive system. Efficient scrolling
requires indepth system knowledge of the CDC-6400 controller,
channel coupler, PDP-11/50, and the COMTAL display system.

The rectangle superimposed on the reduced resolution image of
Figure 3 indicates the fast graphics capability of the system and
here it is used to indicate the area which is being scrolled at full
resolution on the second COMTAL display. Figure 4 is a portion of
this area within the rectangle whereby the interpreter stopped the
scroll and began other operations. In this case, he wishes to

Figure 4.   Scroll/Magnify/Mensurate

extract additional information about a fighter plane parked on the runway. Figure 4 is a four-time, digital enlargement of the airplane, and it is accomplished using a bilinear interpolation process whereby a pixel and its neighbors are used to create additional pixels.

Mensuration is also depicted in Figure 4 in which the interpreter measures the body length and wing/body angle. This is accomplished by successively placing the cursor at both ends of the body and the length is computed almost instantaneously using parameters of the photo such as its scale. Similarly, the angle between the wing and the body is computed by placing the cursor in three successive positions and the result is produced as quickly as the length mensuration. Both length and angle are displayed on the COMTAL screen and these results can be saved by assigning a new name to the image and stored back into the digital data base.

Targeting is depicted in Figure 5. Here, the same image is

Figure 5.   Scroll/Magnify/Target

used by the interpreter but now he wishes to obtain its Universal Transverse Mercator (UTM) grid coordinates. This is a valid, world geodetic coordinate system, and it requires preliminary preparation of the digital photograph in order to place it into this system. We will not discuss this process since it is complex and lengthy. Let it suffice to say that once the data base is prepared, target coordinates are obtained as quickly as mensuration results. The results of a targeting exercise is shown in Figure 5 and the UTM zone number, Northing and Easting and image name are shown on this photo again using the fast graphics capability of the COMTAL system. The coordinates are relative to the center of the circle where the operator placed the cursor.

Once these operations are completed to the satisfaction of

the interpreter, he can continue the scroll until the entire area
within the rectangle has been searched at full resolution. Of
course, there are numerous other functions which he can perform such
as grey-scale remapping as the scroll continues; that is, he can
digitally change the contrast of the photo under cursor control con-
currently with the scroll. This operation requires that both the
PDP-11/50 and CDC-6400 work concurrently and in real time.

Military Intelligence-Digital Filtering. Many times a photo inter-
preter is confronted with suboptimal photography thereby hindering
the interpretation process. Such is the case depicted in Figure 6,
and Figure 7 represents the digitally filtered image. The original

Figure 6. Original   Figure 7. Filtered

image took on a half-tone appearance in the printing process which
masks out details such as the separation between vehicles in parking
lots. These figures are digital enlargements of the pre- and post-
filtering operation and were digitally extracted from their larger
counterparts.

The original enlarged image is 512 x 512 pixels by 8 bits
per pixel, and it was passed through the Fast Fourier Transform (FFT)
software package residing in the parallel processor side of the
interactive system. The FFT is one of those algorithms which is
amenable to being executed in parallel, and the STARAN performs this
operation approximately ten times faster than the sequential pro-
cessor. It is clear that the filtering process removed the half-
tone appearance and that it provides a better chance of interpreting
the contents of the photo.

Mapping Functions -- Profiling Contouring. Line-of-sight problems
are commonplace in a field army environment. Figure 8 represents
discrete elevation data encoded with grey shades, and the profile
between the points connected by the line is shown in Figure 9. In
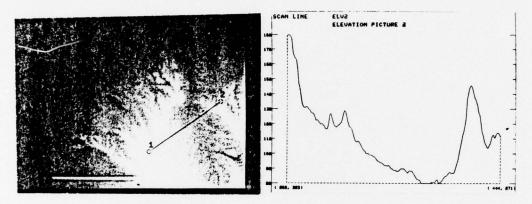

Figure 8.   Gray-Shaded
                  Elevations

Figure 9.   Profile

order to produce the effect shown in Figure 8, 250,000 elevation
points were selected from a section of a topographic map and grey-
shade values assigned to them. The lowest elevation was assigned
black, the highest value was assigned white, and several other grey
values were assigned to those in between elevations, thereby
resulting in pseudo three-dimensional representation of the terrain.

Application software was written to allow the operator to
select the end points of a desired profile by using the cursor. That
profile is then output on a hardcopy device at the operator's
terminal. The profile shown in Figure 9 indicates that an observer
is not able to see from point 1 to point 2. This can almost be
assumed by viewing the grey values at these two points in that point
1 is almost white, while the line to point 2 transverses white areas
also. However, the computer is a better judge of the grey value at
specific points than the human eye, and in addition, the profiling
process is completed in less than 2 seconds.

The digital image created from these elevations can also be
"quick contoured" by assigning colors to bands of grey values
(elevations). The original image is assigned a different color than
those chosen for contours so that they become more vivid on the dis-
play. The result of this process again takes place in real time and
is shown in Figure 10. Although the image is shown here in black

GAMBINO

and white, a vivid contouring operation is displayed on the COMTAL

Figure 10. "Quick" Contouring

screen in pseudocolor. There have been many comments concerning the usefulness of pseudocoloring, and a favorable one is that it indicates subtle changes in the grey shades better than black and white.

Weapon Products -- Pershing Missile. The Army's Pershing missile system is terminally guided to its target by an electro-optical device called the "correlatron." It is housed in the missile along with a Plan Position Indicator (PPI) radar which scans the target area at several elevations during its downward flight. The "live" radar scene is correlated with a prestored, synthetic scene within the correlatron. The match, or mismatch, between these scenes provides the guidance information. Our concern is with the generation of the synthetic scene, and the part played by the digital interactive system.

Synthetic radar scenes are generated from elevation and cultural data bases where slopes derived from the elevations are used to predict radar returns along with returns predicted from the cultural information, such as bridges and buildings. Currently, the initial synthetic scenes are produced in a batch computer environment and the digital, interactive processing begins with this product to make them look even more like the live scene. Live scenes of targets have been digitized, and, as the interactive process progresses, the continually modified synthetic scenes are correlated with live scenes. Correlation curves and statistics are output at the operator's terminal, and the correlation is performed in the parallel processor.

Figure 11 is a synthetic scene produced in the batch computer environment, and Figure 12 is the result of many modifications to the initial synthetic scene. The correlation curves are shown in Figures 13 and 14, where the former indicates poor correlation and the latter good correlation since it has a pronounced peak as opposed to the flatness of the former. The good correlation curve was obtained

Figure 11.   Original
               Synthetic Scene

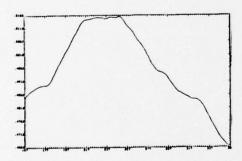Figure 12.   Modified
               Synthetic Scene
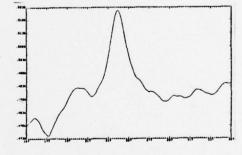

Figure 13.   Poor Correlation

Figure 14.   Good Correlation

after numerous changes to the original synthetic scene, but the correlation was carried out after each change in order to monitor its effect. Correlation is performed by moving the digitized live radar image in a horizontal and vertical fashion over the entire modified synthetic scene, thereby producing 240,000 correlation numbers which are used to produce these curves. This is accomplished in a few seconds on the parallel processor whereas a sequential computer would bog down in this operation. Only a few more seconds are required to display the results and obtain a hardcopy of the curve.

CONCLUSION

A great deal of effort has been expended in the development of the digital interactive system at the U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia. The results of integrating hardware, developing system and application software has been very encouraging since processes which took a week to accomplish in an ordinary computer environment are now taking only a few minutes. This speeds up the R&D process and offers an opportunity to see clearly the advantage of flexibility afforded by digital processing. Nevertheless, some basic problems remain and they will remain for years to come. As stated earlier, a characteristic of digital image processing is data volume. Until experience is gained with processing on conventional computers, it is difficult to assess the impact of digital image processing on these systems. Our experience indicates that it is very easy to inundate the system with digital image processing tasks because of this. This problem will be compounded in the future as acquisition systems of interest to the Army become digital. Also, it is very clear that certain tasks, such as correlation of digital imagery from stereopairs of photographs, is not relegated to the conventional computer because it is too slow for the production of elevation data. These problems, and more, will spur research in digital parallel processing.

We are looking forward to augmenting the existing system with advanced displays and other nonstandard digital hardware and software in an attempt to overcome some of these problems. This approach must be taken in order to demonstrate the practicality, cost effectiveness, and competitive nature of the digital approach relative to other technologies which could be used in performing functions such as those outlined in this paper. Performing the R&D in an ordinary, general purpose computer environment will always leave doubt concerning these attributes.

GAMBINO

REFERENCES

[1] Lawrence A. Gambino and Bryce L. Schrock, "An Experimental
     Digital Interactive Facility," COMPUTER, August, 1977.

[2] Lawrence A. Gambino and Roger L. Boulis, "Staran Complex,
     Defense Mapping Agency — U.S. Army Engineer Topographic
     Laboratories," 1975 Sagamore Computer Conference on
     Parallel Processing, pp. 132-141.

[3] Robert Edwards, Robert Eggert, and John Martin, "Option
     Description, DRHCD, CDC Channel Coupler," Computer Special
     Systems, DEC, California, September 19, 1974.

15

PULSED HOLOGRAPHIC ANALYSIS OF LARGE
VIBRATING VEHICLE COMPONENTS

*GRANT R. GERHART, DR.
GREGORY ARUTUNIAN, MR.
TANK-AUTOMOTIVE RESEARCH AND DEVELOPMENT COMMAND
WARREN, MI 48090

INTRODUCTION

Double-pulsed holography provides a measurement technique
for the analysis of large surface areas of vibrating components. A
particular advantage of holography is that it requires no contact
with the vibrating object, and it gives a solution to a classical
problem in the area of both acoustic and vibration analysis. The
surface of a vibrating object is represented by a series of fringes
which connect points that have equal amplitudes of displacement.
Since holography is based on interferometry, the accuracy is within
a fraction of the wavelength of light.

Continuous wave (CW) holographic techniques, which utilize
HeNe or argon laser light sources, are limited in a number of ways.
The test object in CW holography has to be isolated from the environ-
ment because the maximum tolerable motion of the experimental appar-
atus is about an eighth of the laser light wavelength. CW holographic
interferometric techniques are time-average and real-time holography,
and they are restricted to the measurement of very small amplitudes.
A maximum amplitude of displacement of about 25 wavelengths is
possible using strobe light techniques such as acoustical optical
modulators.

The difficulties in CW holography can be circumvented by
the use of double pulsed holographic techniques. This work is based
on results obtained by a Korad holographic camera, which employs a
Q-switched ruby laser. Each laser pulse has a duration of twenty

nanoseconds which is insignificant compared to the periods of vibrating structures. The pulse separation time usually varies between a few microseconds and about one millisecond while typical values usually range between 200 and 500 microseconds. Pulse separation time intervals are kept below one millisecond to avoid isolation problems with low frequency building vibrations but are still long enough to record surface displacement occuring at the frequencies of interest. In any case the pulse separation times are much shorter than the periods of the vibration frequencies for large amplitude displacement.

PULSED RUBY LASER

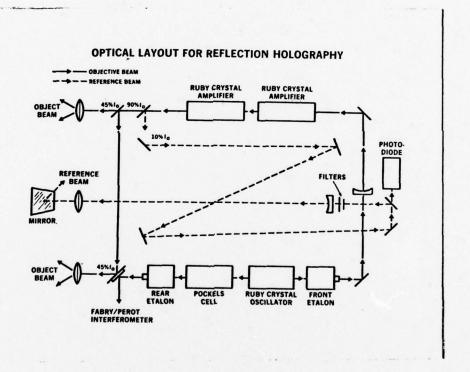**OPTICAL LAYOUT FOR REFLECTION HOLOGRAPHY**

Figure 1

Figure 1 shows a diagram of the pulsed holographic camera which is used for our data acquisition. The laser oscillator is located at the bottom of Fig. 1 and it produces a low energy light pulse of a few millijoules. The oscillator cavity consists of a front and rear etalon, pockels cell, and ruby rod. The two etalons each contain some Fabry-Perot elements and are individually tuned by controlling the temperature to within a tenth of a centigrade degree. Individual resistive heating elements with the appropriate control circuits are used to maintain a very uniform etalon temperature over relatively long time intervals. In addition the etalons are in protective housings which will insulate them from the ambient temperature. The pockels cell is used to Q switch the oscillator and provide two output pulses of equal amplitude. The etalons are tuned so that the ratio of front-to-rear output energy is a maximum. The output from the rear etalon is diverted into the Fabry-Perot interferometer. The ruby rod is also temperature controlled by a separate water cooled unit to within 0.1C.$^{\circ}$ Special anti-reflecting coatings are on the rod surfaces so that the cavity length is determined by the two etalons. The rear etalon has four elements, and it is designed so that the oscillator cavity modes have a high finesse and a very narrow line width for long temporal coherence lengths. The front etalon has a single quartz element and is tuned for maximum output.

The output from the oscillator is diverged by a negative lens before entering the first of two ruby amplifiers. The output from the ruby power amplifiers consists of two pulses with up to two joules per pulse that passes through a beam splitter. Ten percent of the light energy is reflected into the reference beam while the remaining 90% is equally distributed between the two object beams. These beams then pass through diffusers in order to obtain a uniform illumination of the object. The amplifier output beam energy is monitored by sampling a part of the reference beam with a calibrated photo-diode. The reference beam passes through a two meter delay line and uniformly illuminates the holographic plate.

Figure 2 shows the holographic camera head. The circular window in the side cover allows the beam from the rear etalon to pass through to the Fabry-Perot interferometer which is not present in this photograph. The plate holder is shown at the top with an electronic shutter which is synchronized with the pulse trigger. The two object ports are shown in front along with the reference beam mirror housing.

Figure 3 shows a side view of the laser head with the Fabry-Perot interferometer mounted on the left side. The ruby

oscillator is clearly displayed along with the high voltage cables
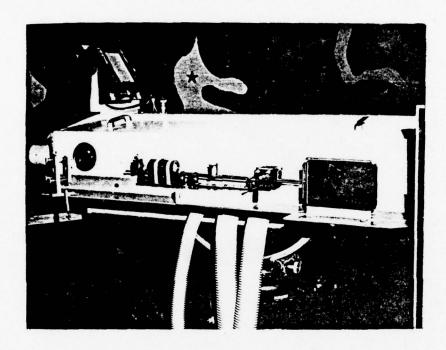and cooling lines.

Figure 2

Figure 4 contains two instrumentation racks which house
various power supplies and pulse shaping components. The left rack
contains two 10KV power supplies for the ruby amplifiers and the
oscillator flash lamps. The right rack also contains two 10KV
power supplies for pulsing each side of the pockels cell to obtain
the double pulse output. An additional component in this rack is a
digital pulse generator which sets the pulse separation time between
each of the output pulses. Also in Figure 4 are a storage oscillo-
scope for displaying the photo diode outputs and two digital
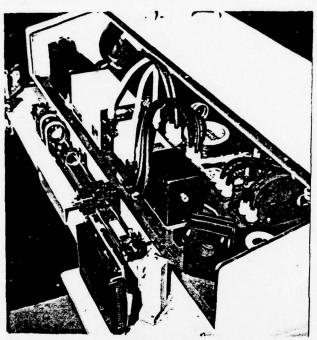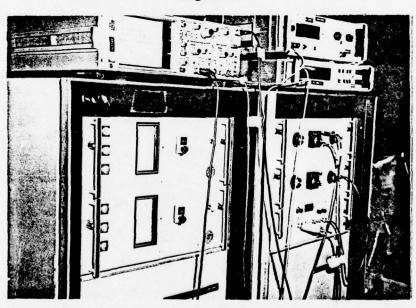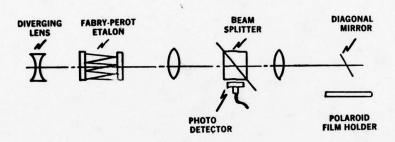thermometers for monitoring the temperature of the etalons and the
ruby rod.

Figure 3

Figure 4

FABRY-PEROT INTERFEROMETER

DIVERGING LENS   FABRY-PEROT ETALON   BEAM SPLITTER   DIAGONAL MIRROR

PHOTO DETECTOR   POLAROID FILM HOLDER

Figure 5

Each point of the illuminated object reflects a portion of the laser light back onto a photographic plate as shown in Fig. 1. Since the object and reference beams are coherent, an interference pattern is created on the photographic plate. A hologram is a photographic place which contains both the amplitude and phase information of the reflective wave from the object. A virtual image of the object is then reconstructed by illuminating the hologram with light from a HeNe gas laser.

In double pulse holography two consecutive exposures of a vibrating object are made on the same photographic plate. These exposures are coincident with the double pulse output of the ruby laser and record information about the position of the vibrating

object at the two distinct times. The interference between these diffraction patterns during reconstruction gives rise to the fringe pattern.

RUBY LASER COHERENCE REQUIREMENTS

The ruby laser must have a coherence length of at least one meter and preferably closer to 7 meters. The coherence limitation of our system is temporal and not spatial. A temporal coherence of 7 meters for a double pulse laser implies that both pulses must oscillate on the same cavity mode. The coherence length is reduced to one meter if the laser pulses are of equal intensity and located on adjacent cavity modes. The laser cavity contains components for longitudinal and transverse mode selection to achieve long coherence-length pulses. The rear etalons (Fig. 1) contain numerous elements to suppress longitudinal mode components. The output energy from the ruby laser oscillator is drastically reduced after mode selection. The laser is tuned to oscillate in the TEM_{oo} mode just above an energy threshold of about 5 millijoules per pulse. The laser output is Q-switched to produce a pulse width of 20 to 30 nanoseconds. The peak power from the double stage amplifier output is 100 megawatts.

The coherence length between each of the two pulses can be monitored using a Fabry-Perot (FP) interferometer (Fig. 5). The FP unit is mounted on a bracket alongside of the holographic camera as shown in Fig. 3. The FP interferometer consists of a FP etalon, a diverging lens, and a Galilean telescope with a magnification of about 15. A Polaroid camera back records the FP interference pattern which consists of circular rings of various orders. The laser output is multi-moded if the various interference orders have a multiple structure. Fig. 6a shows a FP interference pattern for a double pulse, single mode laser output. Each of the diffraction orders is a singlet and the relative temporal coherence of the two pulses is greater than 7 meters. Figure 6b shows a FP interference pattern where the output energy is distributed over several cavity modes. The temporal coherence of this double pulse shot is only a few centimeters.

A great deal of operator skill and tuning is required to maintain a 7 meter coherence length for double pulse operation. Temperature effects are very important and a temperature controlled laboratory is a must for consistent results.
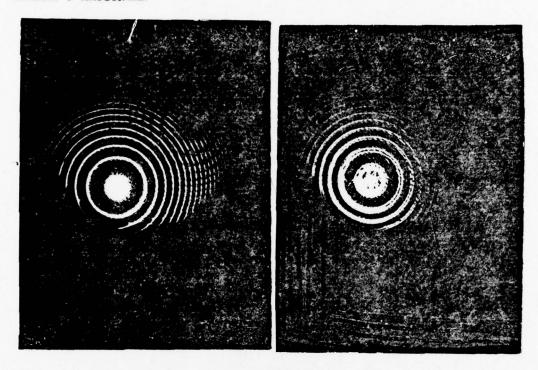
Figure 6a   Figure 6b

A CALCULATION OF THE VIBRATION AMPLITUDE

Consider a point on the surface of a vibrating object which is illuminated by two pulses from a ruby laser. The pulses are of equal intensity and separated by a time interval which is small compared to the period of oscillation T. The point on the surface moves a distance d during T and generates a phase difference ϕ between the two object points as shown in Fig. 7. The optical path difference between the two object waves is given by the amount

$$d = d(\cos\alpha + \cos\beta), \tag{1}$$

where α and β are the angles of illumination and reflection with respect to the surface normal. The phase difference ϕ is

$$\phi = \frac{2\pi d}{\lambda} \ (\cos\alpha + \cos\beta), \tag{2}$$

where λ is the wavelength of the ruby laser. Each object wave inter-feres with the reference beam to produce a separate hologram on the

photographic plate.

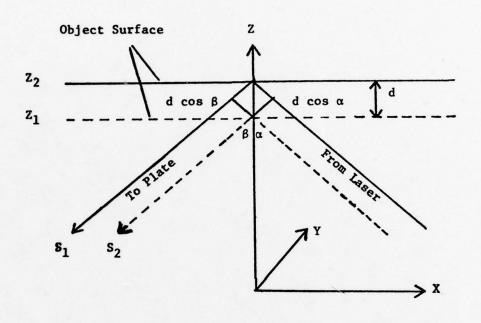

Figure 7

The superposition of the two exposures during reconstruction yields a series of fringe lines. A minimum in intensity is obtained if

$$d_{min} = \frac{(2n + 1)\lambda}{2(\cos\alpha + \cos\beta)} \qquad (3)$$

where n is a positive integer. Points of equal displacement are connected by interference fringes, and the relative deformation of the object surface is determined by the contour and spacing of the holographic fringes. The spacing between n interference fringes gives the component of the displacement normal to the surface as

$$d_{max} = \frac{n\lambda}{(\cos\alpha + \cos\beta)}. \qquad (4)$$

The relative displacement between the two object points is determined by counting the interference fringes between these points.

The average surface velocity is the relative displacement divided by the pulse separation time Δt, and it is proportional to the density of holographic fringes.

A double exposure hologram is used to measure the amplitudes of vibration. A transducer which is attached to the area of interest provides the trigger signal for the laser pulses. Any part of the period of oscillation can be examined by employing electronics delay circuitry. The relative timing of the laser pulses and the oscillation of the object point is shown in Fig. 8.
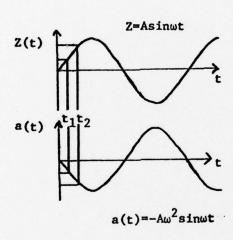

Figure 8

The displacement $Z(t)$ of the object surface at times t_1 and t_2 is given by

$$Z(t_1) = A \sin \omega t_1$$

$$Z(t_2) = A \sin \omega t_2$$

$$d = Z(t_2) - Z(t_1), \qquad (5)$$

where ω is the oscillation frequency of the object point. The amplitude A is determined by substituting Eq 5 into Eq 3. Amplitudes of any magnitude can be determined at the object point by using this technique.

DISCUSSION OF EXPERIMENTAL RESULTS

A detailed analysis of three holograms will be discussed in this section. A 15 milliwatt helium-neon laser was used to reconstruct the virtual holographic images which were quite bright and easily viewed with the naked eye. These images were photographed with Polaroid 4 by 5 inch type Positive/Negative film. A 105 mm telephoto lens with a F16 stop gave quite good results. The contrast ratio of the holographic fringes was better than the photographic film; consequently, the photographs are not as good as the holographic images. The photographs are also inferior because of their limited depth of field and field of view.
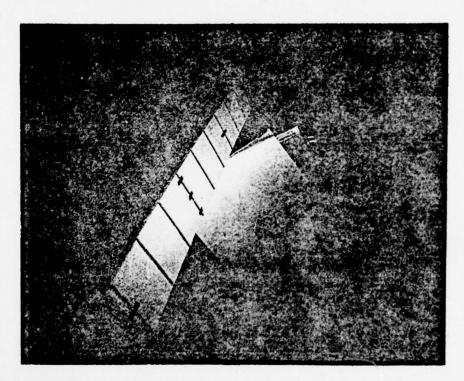

Figure 9

Figure 9 is a photograph of a single pulse, single mode hologram with a 3.5 meter calibration board extending across the field of view. The black lines on the board are approximately 20 cm. apart and about 75% of the board is visible in the photograph. The entire board is visible in the actual holographic image in addition to the wall in the background. The hologram has at least seven meters coherence properties since no localized fringes are visible along the entire extent of the board.
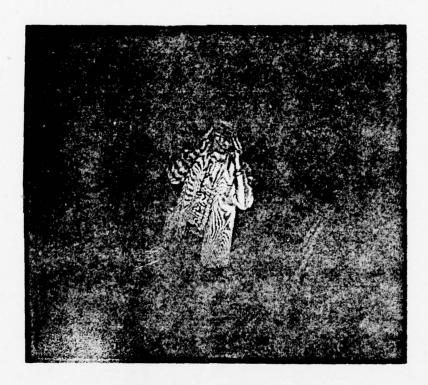

Figure 10

Figure 10 is a photograph of a seven meter double pulse hologram of a man standing 2 meters in front of the holographic camera. The man is apparently motionless to the unaided eye; although, many fringes exist in the holographic images. The tiny muscle movements of the human body are clearly visible via the fringe contours, and the regions of highest fringe density are regions of largest amplitude of vibration. The fringe contrast in this hologram is excellent because the two ruby pulses have nearly identical amplitudes. The pulse separation time between the two laser pulses is

about 500 μ sec.

Figure 11

Figure 11 is a photograph of a double pulse hologram of a vacuum pump. The fringes are caused by an internal vibration due to the motor running at a normal rate. The pump is painted white to increase the fringe contrast and the pulse separation time is 300 μ sec. The parallel fringes on the baffle and the intake tube are caused by a cantilevered motion of these parts. The bulls eye patterns on the base and the side panels are due to mode excitations on these respective regions of the pump.

None of the double pulse holograms were synchronized with the vibrational motion of the object. The ruby laser was manually triggered so that the fringe pattern varied depending upon the relative phase between the trigger pulse and the object motion. The fringe contours remained relatively constant in shape; although, the fringe intensity at a fixed point on the object varied randomly between light and dark. In the future we intend to externally

trigger the laser off of an accelerometer transducer which is fixed
on the object. The fringe contour and spacing will remain constant
if the system parameters are unchanged from one hologram to the next.
This procedure will help standardize the variation in fringe pattern
which is caused by real structural changes in the object.

The holographic camera in its present configuration is a
very temperamental piece of instrumentation. The main problem seems
to be one of temperature stabilization. If the laboratory temperature
varies by more than 5 Celsius degrees, the laser output becomes
unstable. Since the oscillator must operate so close to threshold,
small variations in cavity length or element spacing in the etalons
rapidly detunes the system. The etalon temperatures seem to be well
controlled by resistive heaters which are positioned on the mirror
mounts. The most likely problem at this time is the temperature
variations in the aluminum baseplate. Work is in progress at Korad
to mount the oscillator on an invar base plate. Since the oscillator
components are bolted directly to the base plate, the much smaller
thermal expension coefficient of invar should minimize this effect.
Another approach is to put resistive heaters on the current aluminum
base plate and heat it above the ambient. The base plate will be
insulated from the ambient and its temperature will be controlled to
within 1 centigrade degree. A last resort would be to put the
oscillator in a precisely temperature controlled oven and insulate
it from the ambient. This approach would be quite expensive and
involve a considerable redesign of the present holographic system.

Pulsed holography promises to be a very exciting and useful
tool for investigating the vibration characteristics of mechanical
structures. The fringe pattern gives the design engineer a visual
picture of regions on the structure with large displacement ampli-
tudes, large stress concentrations, and the position of nodes and
anti-nodes of vibration. The state-of-the-art ruby lasers have
large energy outputs that enable the holographer to illuminate sizable
surface areas on the test specimen. The basic limitation is that at
present the technique is suitable only for laboratory environments.
The tight temperature specifications of the oscillator would make
routine double pulse holography difficult in a production line
environment.

NOVEL CONCEPTS IN REAL-TIME OPTICAL TRACKING

ALTON L. GILBERT, DR.
*MICHAEL K. GILES, DR.
US Army White Sands Missile Range
White Sands Missile Range, New Mexico 88002

INTRODUCTION: Optical tracking has been a mainstay of accurate metric range instrumentation since the first testing of modern rocketry. The accuracies that were possible from optical instruments exceeded those from other available instruments. Improvements in encoders, optical testing, modelling of the atmosphere, and optical design continuously improved the accuracies of optical instruments. The major drawback is the required film processing which delayed the delivery of boresight corrected optical data.

Recent changes in technology have created the potential for relieving part of the delay in data delivery. Automatic tracking methods using high-speed microprocessors, artificial intelligence, and pattern recognition techniques, together with special modifications to the existing optical systems, are now available to perform most of the film reading function in an on-line, real-time mode. These methods far exceed the conventional contrast, edge, and correlation trackers in sophistication and capability, since they are based upon an understanding of some definable properties of the image involving many parameters as compared to only a few.

THE INTELLIGENCE OF OBJECT IDENTIFICATION: Pattern recognition is a mathematical science based upon the separation of a parameter space into two or more regions, so that when the parameter is measured it may be classified as belonging to one of the appropriate regions. It follows that a vector parameter will give rise to a parameter space of dimensionality equal to the number of independent elements in the vector. Thus, for an N-vector, the required separation is a hyperplane in N space. If the parameter is a single element vector, an assignment

1

.can be made on the basis of a single threshold on the real numbers, and a tracker can be built that uses this decision rule. An example we call a contrast tracker uses a threshold on brightness for the assignment. A preprocessing algorithm may be placed before the decision. If we pre-process for the magnitude of change in intensity, the same thresholding rule will yield an edge tracker. These are amongst the simplest appli-cations of pattern recognition to the object identification problem. Since these algorithms are easily confused, many spurious objects in the field of view (FOV) often meet the classification criteria.

A somewhat different approach that uses an array of points and measures the closeness of fit to a subsequently measured similar array, while choosing the best match as the correct location, is generally known as a correlation tracker. The decision is again based upon a sin-gle element parameter vector (the closeness of fit), but the preprocess-ing of data is much more elaborate. While this approach offers an improvement in confidence that the correct object has been located if the object description is known, it suffers from two principal problems. The first is that generally the object being tracked changes appearance continually while objects in the background may not. This requires an adaptive object description which may slowly converge to the acceptance of an undesired object as the desired one. The second problem is that this approach requires a very large amount of processing to do a good job, since the optimal linear process would be a convolution of the NxM object description array over the PxQ array of data points, and gener-ally P>>N, Q>>M. A commonly used simplification is an additive (sub-tractive) algorithm that seeks the best fit of the desired array to the data, instead of the convolution. This approach necessarily results in loss of tracker performance. For these reasons, the correlation track-ing method is generally limited to very restricted window tracking and fairly slow update rates.

Approaches to real-time optical tracking have generally been limited to these approaches for the following principal reasons. The first and most important has been the magnitude of the real-time proc-essing requirement for the more elaborate approaches, which have exceed-ed computational resources generally available. The second has been a lack of image understanding that would allow the formulation of more reliable, yet simple, approaches. Substantial progress has recently been made in the former, and there are many encouraging new developments in the latter.

A variety of methods of image data processing have become known over the past decade. Applications-oriented research at the US Army White Sands Missile Range (WSMR) has lead recently to a system of reasonably high sophistication using concepts developed in-house and

through sponsored research to solve complex identification and track-
ing problems. WSMR has concentrated on objects in the visible spectrum
and in real-time. Many other systems, not necessarily real-time, have
been developed for applications in medicine, meteorology, and space
research.

Many of the newer methods involve the use of many elements in
the parameter vector to glean more information from the data. In apply-
ing pattern recognition methods to the object identification problem,
the engineer is trying to minimize the amount of data he must handle and
maximize his confidence that he made the correct decision. Any linear
process will preserve the quantity of data (260,000 points for a 512x512
image, possibly 8 bits per point) which is obviously not desirable if
much processing is required to make a decision. The engineer is forced
to require a high degree of parallel processing on linear processes, and
to perform nonlinear operations to reduce the data quantity prior to
determining the values of the parameter elements used in the decision
rule. Ideally, the dimensionality of the decision space should be kept
reasonably small to allow decisions to be made in real-time or in near
real-time.

Some of the preprocessing methods currently in use are:

Filtering: Filtering operations generally involve the con-
volution of a point spread function array with the image to achieve some
desired objective with the image. Examples include removing spatially
invariant degradations due to the optics of the atmosphere, boosting the
high frequency content of the image to enhance edges, removing noise in
the image, making the image more pleasing to the eye, and other such
operations. Generally those operations which remove degradations are
called estimation and those which emphasize certain spatial frequencies
or certain aspects of the image are called enhancement. It must be
noted that enhancement is an intentionally introduced distortion to pro-
duce some desired effect.

Transforms: Operations that map the image into a new domain
are called transforms. The elements in the new domain are a measure of
some property of the original image. The most common example is the
discrete Fourier transform (DFT), especially in the fast algorithms
(FFT). The DFT identifies the spatial frequency content of the image,
which allows further processing based upon these components. A class
of binary Fourier (BIFORE) transforms has been developed over the past
decade which are similar to the DFT but are more suited to computer
applications. These might be called lesser transforms since they do
not represent the information in the image as completely. Because they
are much more efficiently run on a computer than the DFT, they have

important applications in image transformations. Among these lesser transforms are the now popular Hadamard transform based upon Walsh functions, and the less known but simple Haar transform. These transforms can be useful for identifying features of interest in the image. It is necessary, of course, to apply all of these transforms in a two-dimensional algorithm to process the two-dimensional images.

Point Processing: In point processing, individual points in the image are assigned new values based upon some assignment rule. This may take a variety of forms with a large variation in apparent results. One point processing algorithm averages the corresponding point of several frames or sequential images to produce a weighted composite and remove transient degradations. Another assigns all values above a given threshold to 1 and all values below to 0. This is known as thresholding. A variation on thresholding is to assign predetermined gray levels to 1 even though these may not be in a continuous range. Still another algorithm, known as contrast stretching, assigns all values below some intensity I_0 to 0; all values above another intensity I_1 to the maximum gray level, say 256; and stretches the intermediate values to occupy the full range. Generally, point processing methods are nonlinear, yielding fewer bits in the output than in the data array.

The next step in the process is to identify the values of the elements in the parameter vector. These elements may include such things as size, orientation, number of corners, brightness, etc. When joined in a single parameter vector, they describe all we think we need to know to adequately describe the object for purposes of identification.

A REAL-TIME TRACKING SYSTEM: By using the above concepts together with high-speed microprocessors and special optics, a real-time tracking system may be devised that demonstrates a substantial advantage over the contrast, edge, and correlation trackers currently on the market. The greatest challenge is that of doing "intelligent" processing of video data at the extremely high data rates of standard TV.

The development of an intelligent real-time video (RTV) tracking system has been accomplished through the cooperative efforts of research and development personnel at WSMR, New Mexico State University (NMSU), and the Optical Sciences Center of the University of Arizona. The prototype RTV processor is being assembled at NMSU, the automatic zoom lens and image rotator at the University of Arizona, and the system interfaces at WSMR. The system components will be integrated and the system deployed early in fiscal year 1979 as an add-on modification to the Contraves Model F cinetheodolite at WSMR.

Figure 1 is a block diagram of the RTV tracking system which shows the RTV processor as the central element. The RTV processor
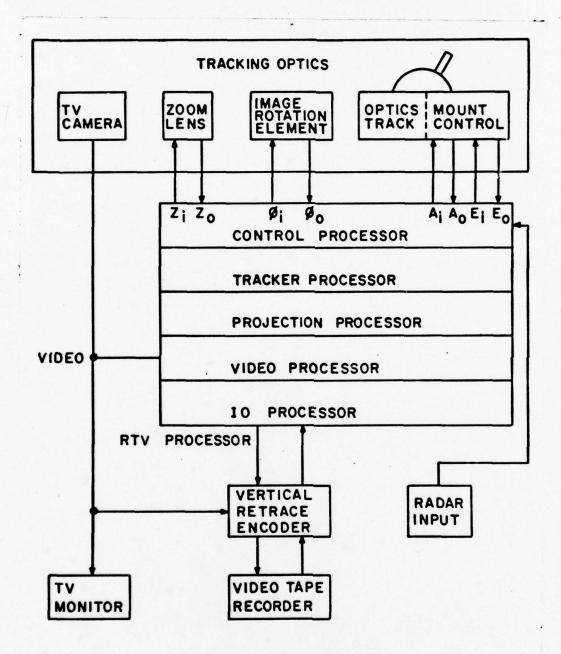
4

TRACKING OPTICS

TV CAMERA

ZOOM LENS

IMAGE ROTATION ELEMENT

OPTICS TRACK | MOUNT CONTROL

Z_i Z_0 $\emptyset_i$ $\emptyset_0$ A_i A_0 E_i E_0

CONTROL PROCESSOR

TRACKER PROCESSOR

PROJECTION PROCESSOR

VIDEO

VIDEO PROCESSOR

IO PROCESSOR

RTV PROCESSOR

VERTICAL RETRACE ENCODER

RADAR INPUT

TV MONITOR

VIDEO TAPE RECORDER

FIGURE 1. RTV TRACKING SYSTEM

5

receives standard composite video from a television camera, locates the target image, and provides control signals which drive the zoom and image rotation elements and point the Contraves tracking optics at the target. It also provides boresight correction signals and target attitude angles which are recorded into the vertical retrace period of the video tape used to record the tracking sequence.

The RTV processor consists of a distributive array of five processors, shown in Figure 1. The video processor synchronizes and digitizes the video signal from the TV camera, performs a statistical analysis of the digitized image, and separates the target images from the background. The projection processor accumulates binary projections of the target and plume images and establishes the structural parameters which locate and describe the shape of the target and plume images. The tracker processor establishes a structural confidence in the data and implements an intelligent tracking strategy. The control processor utilizes the structural confidence to combine current target coordinates with previous target coordinates to orient the optics toward the next expected target position, forming a fully automatic system. The input/output (I/O) processor provides a user interface to the tracking processors and is responsible for recording the tracking data with a video tape recorder.

A Research Oriented Processor Configuration: Four of the five distributive processors (excluding the I/O processor) which comprise the RTV processor are high-speed microprogrammable processors, each of which requires a stored microprogram to control its designated tracking function. To provide a powerful tool for future research in video tracking algorithms and to facilitate operational testing of the RTV system, the control store of each processor is realized with a read/write random access memory.

These four distributive processors are being built with a standard microprogrammable processor architecture to simplify the development and maintenance of the RTV tracking system. This standard architecture has been designed, built, and tested at NMSU. Based on the new Texas Instruments (TI) 74S481 Schottky processor chip, it provides a microinstruction cycle time of under 200 nanoseconds with sufficient computational power to implement the required RTV tracking algorithms. The standard architecture requires several LSI chips which may be partitioned into control and processing sections. Overlapping the execution of one microinstruction with the fetch of the next one allows the processor to achieve a minimum microinstruction cycle time equal to the larger of either the fetch time or the execution time, significantly increasing the speed of the processor.

The four high-speed processors included in the RTV tracking loop are described in some detail in the following paragraphs. In each case, the processor is built around the standard architecture outlined above. Some specialized hardware is added to the standard configuration in each case to accommodate the specific functions of the individual processors.

The Video Processor: The video processor decomposes each video field into target, plume, and background pixels at the standard video rate of 60 fields per second. As the TV camera scans the scene, the video intensity is digitized at m equally spaced points across each horizontal scan line. A resolution of m = 512 pixels per line results in a pixel rate of 96 nanoseconds per pixel. Within 96 nanoseconds, a pixel intensity is digitized and quantized into 8 bits (256 gray levels), counted into one of six 256-level histogram memories, and then converted by a decision memory to a 2-bit code indicating its classification (target, plume, or background). The 2-bit classification code is passed to the projection processor via the target data (TD) and projection data (PD) lines. TD is high for target points; PD is high for plume points.

The basic assumption of the image decomposition method is that the target image has some video intensities not contained in the immediate background. A tracking window is placed about the target image, as shown in Figure 2, to sample the background intensities immediately adjacent to the target image. The window frame is partitioned into two regions, B and P. Region B is used to provide a sample of the background intensities, and region P is used to sample the plume intensities when a plume is present. Using the sampled intensities, a very simple decision rule is used to classify the pixels in region T as follows:

- Background points--All pixels in region T with intensities found in region B are classified as background points.

- Plume points--All pixels in region T with intensities found in region P, but not found in region B, are classified as plume points.

- Target points--All pixels in region T with intensities not found in either region B or P are classified as target points.

A tracking window placed about the target image provides a method for sampling the pixel features associated with the target and background images. The background sample should be taken relatively close to the target image, and it must be of sufficient size to accurately characterize the background intensity distribution in the vicinity of the target. The tracking window also serves as a bandpass filter
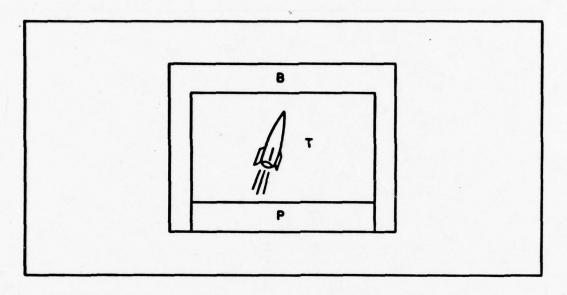
7

FIGURE 2. TRACKING WINDOW

by restricting the target search region to the immediate vicinity of the target. Although one tracking window is satisfactory for tracking missile targets with plumes, two windows provide additional reliability and flexibility for independently tracking a target and plume, or two targets. Having two independent windows allows each to be optimally configured and provides reliable tracking when either window can track.

The Projection Processor: The projection processor consists of a projection accumulation memory (PAM) and a standard processor which are designed to form projections of simultaneous target and plume windows and to compute structural parameters from the projections. The pixel data from each tracking window enters the PAM in real-time as a synchronized serial stream on lines TD and PD. As the classified pixel data is received, the PAM accumulates the projection data while the processor monitors the y-projections, accumulates the total number of target and plume points, and determines the midpoints used to split the x-projections. Each x-projection is split to allow the computation of target and plume attitude angles based on the locations of the median centers of the x- and y-projections of the top half and bottom half of the target and plume images.

During the vertical retrace interval, the projection processor divides each projection into eight segments of equal mass using a simple algorithm to sequentially address each line of the projection and multiply the number of pixels in the line by eight. If the result exceeds the total number of pixels in the projection, a flag is sent to the PAM forcing the next line to be placed at the beginning of the next 1/8 segment of the projection. If the result is less than the total number of pixels in the projection, additional lines of pixels are accumulated until the line containing the 1/8 percentile point is located.

The 1/8 percentile points for each of the six projections are computed within 410 μsec of the vertical retrace period and then passed to the communication memory along with the total number of target and plume points. These parameters constitute the structural parameters used by the tracker processor to define an intelligent tracking strategy. Figure 3 illustrates the accumulation of the projections and the computation of the percentile points and, for simplicity, omits the splitting of the x-projection.

Tracker Processor: The tracker processor receives the structural parameters from the projection processor, locates and characterizes the structure of the target and plume images, and decides on a tracking strategy to maintain track. It then outputs control signals to place the window frames in the video processor and outputs target location and orientation data to the control processor along with a confidence in the measured data. Since it operates on the projection data from field n while the projections for the next field (n+1) are being accumulated, the tracker processor is always one field behind the video and projection processors. The tracker and control processors must both finish their calculations before the vertical retrace interval begins for field n+1. This constraint requires the tracker processor to output its data to the control processor within 7 milliseconds after it receives the projection data.

Since the tracker processor is the only processor that communicates with all of the other three processors, each of which has its own coordinate system, the tracker processor must interpret the input data intelligently and then output the appropriate data to the video and control processors in their respective coordinate systems. The inputs are positive 16-bit integers defined for a coordinate system whose origin is the first pixel scanned inside the appropriate tracking window. The outputs to the video processor are 9-bit positive integers defined for a coordinate system whose origin is the first pixel scanned within the FOV. The 16-bit outputs to the control processor are defined for a coordinate system whose origin is the boresight.
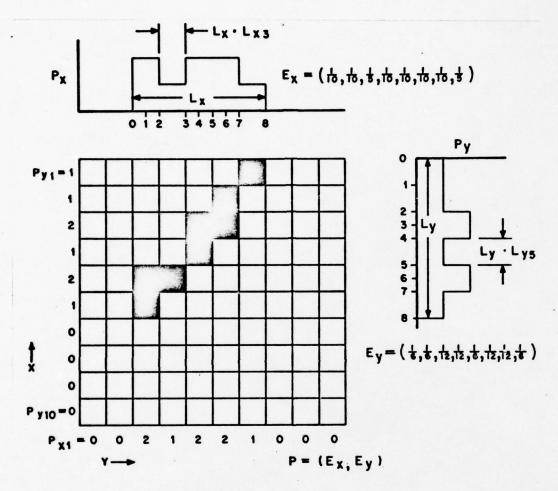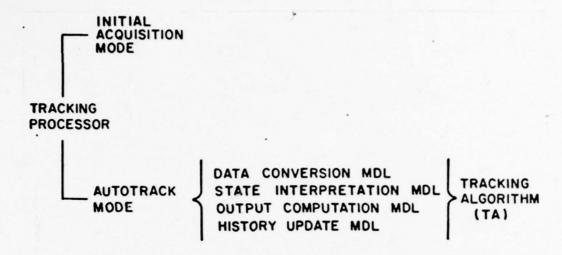
FIGURE 3

PROJECTIONS AND PERCENTILE POINTS

An overall view of the functions of the tracker processor is given in Figure 4. It has two modes of operation, the initial acquisition mode and the autotrack mode. The initial acquisition mode is used when the RTV system is trying to lock onto the target of interest. During this mode, the video processor does little or no learning on the target and plume intensities. The tracker processor will not instruct the control processor to begin predicting the target location until it is sure of the existence of at least the plume within the plume window.

10

```
         ┌─ INITIAL
         │  ACQUISITION
         │  MODE
         │
TRACKING │
PROCESSOR│
         │
         │              DATA CONVERSION MDL  ⎫
         │  AUTOTRACK   STATE INTERPRETATION MDL ⎪  TRACKING
         └─ MODE        OUTPUT COMPUTATION MDL ⎬  ALGORITHM
                        HISTORY UPDATE MDL  ⎭   (TA)
```

FIGURE 4. TRACKER PROCESSOR FUNCTIONS

When the plume image moves into an appropriate region of the FOV, the tracker processor will notify both the video processor and the control processor with a flag indicating that it is now ready to shift into the autotrack mode.

The autotrack algorithm is divided into the four main modules shown in Figure 4. The data conversion module transforms the projection input data into physical variables; such as, target and plume size, position, and shape. These variables are then combined with previous target activity data from the history update module to obtain additional variables; such as, the changes in target and plume position and size. All of these variables are compared with preassigned reference constants to obtain a set of binary inputs which are used directly by the state interpretation module to define the current tracking situation and produce an optimum tracking strategy. The strategy is implemented by the output computation module in the form of control signals to the video and control processors.

The Control Processor: The function of the control processor is to generate the four control signals that drive the real-time video tracker; i.e., the tracker azimuth A_i and elevation E_i which are sent to the RTV-Contraves system interface and the optics rotation ϕ_i and zoom Z_i which are sent to the RTV-zoom/rotation interface (Figure 1). In addition, the control processor outputs the following tracking data to

the I/O processor after each field so they can be recorded in the vertical retrace period of the video tape: field count, tracker status, time, x-displacement from boresight, y-displacement from boresight, tangent of the target orientation angle from vertical boresight, target azimuth, target elevation, tracker azimuth, tracker elevation, image rotation angle, and zoom ratio.

The tracking optics feeds the target image to the video processor portion of the RTV processor (Figure 1) which establishes the target coordinates with respect to the optics boresight. The control processor combines current target coordinates with previous target coordinates to point the optics toward the next expected target position. The predicted control equations are based on the combination of linear and quadratic optical estimates taken from a five-deep history stack. Since the input data is derived from field (K-1), and the estimates are being computed during field K, the control estimates must predict ahead two time increments to provide control signals which will place the boresight at the correct position during frame K+1.

COMPUTER SIMULATION OF THE REAL-TIME VIDEO TRACKER. A computer simulation of the RTV tracking system, incorporating the algorithms used in the control stores of the four distributive processors, has been developed and implemented on the PDP 11/35 system at WSMR. The purpose of this simulation is to provide a method for testing new design concepts and evaluating the RTV tracking system under realistic tracking conditions. The simulation model includes dynamic models for the target trajectory and the Contraves Model F cinetheodolite tracking system, in addition to the RTV processor algorithms, for simulating the complete tracking system. The recent development of an image processing laboratory at WSMR has enabled research personnel to digitize sequential video fields of typical tracking imagery. These fields of digitized video are now being used in the RTV simulation and in the development of improved image segmentation and structural analysis algorithms.

The RTV simulation is being used as a research tool at WSMR. It is especially effective in evaluating the RTV system performance and in identifying and seeking solutions to real-time tracking problems before the RTV tracking system is deployed. With the added capability of using digitized video from a variety of tracking sequences as inputs to the video processor, the simulation can now test the system performance under a variety of tracking conditions, thus allowing thorough evaluation and possible refinement of the tracking and processing algorithms and the state transitions of the tracker processor.

CONCLUSION: RTV tracking is not new, but recent developments have added new capabilities that enhance the advantages of these systems. Video tracking offers some distinct advantages over electronic

12

tracking (such as ECM immunity), but suffers from some disadvantages as well (such as restrictions in visibility). Several other aspects of system development for RTV tracking are discussed in the papers and reports listed in the bibliography.

A continuing research need exists for better understanding of imagery. A human incorporates many elements into the parameter vector that he uses to identify an object. The difficulty of understanding the human visual process has caused rather slow progress in teaching computers to "see." We know that the human uses such things as texture, orientation, color, size, shading, shape, context, etc. to identify objects. Parameter vectors which incorporate these elements are difficult to quantify. It is not necessary, however, to require the computer to see the same things a human does. It is difficult to visualize elements of parameter vectors that do not have a physical meaning to a human, but which may be useful for computer recognition processes. Much work remains to be done to produce a highly sophisticated sight process in a computer.

The concepts described in this paper have, however, been tested and will result in a prototype system deployed in 1979. Through a process of simulation and breadboard verification, WSMR has determined that such a system is well within the current capabilities of technology. A great deal of national (and some international) attention has been focused on this project because of the unique applications of pattern recognition in a tracking situation.

CHRONOLOGICAL BIBLIOGRAPHY: The papers and reports listed below have been produced during the course of this project.

1. Flachs, G. M., "A New Target Tracking Algorithm for the Automatic Programmable Film Reader," WSMR Technical Report, STEWS-ID-74-3, October 1974.

2. Flachs, G. M. and A. L. Gilbert, "Automatic Boresight Correction in Optical Tracking," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, June 1975.

3. Flachs, G. M., "Real-Time Tracking Algorithm," Final Report for Grant DAHCOA-15-G-0038, Army Research Office, October 1975.

4. U, Yee Hsun and G. M. Flachs, "Structural Feature Extraction by Projections," Proceedings of the IEEE Region V Conference, 1976.

5. Black, R. J., J. T. Whitney, G. M. Flachs, and A. L. Gilbert, "A Pre-Prototype Real-Time Video Tracking System," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, May 1976.

13

6. Flachs, G. M., W. E. Thompson, Yee Hsun U, and A. L. Gilbert, "A Real-Time Structural Tracking Algorithm," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, May 1976.

7. Thompson, W. E. and G. M. Flachs, "A Structure and Dynamic Mathematical Model of a Real-Time Video Tracking System," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, May 1976.

8. Flachs, G. M. and J. A. Vilela, "A Structural Model for Polygon Patterns," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, May 1976.

9. Gilbert, A. L. and G. M. Flachs, "A New Concept in Optical Tracking using Pattern Recognition," International Symposium on Information Theory, June 1976.

10. Vilela, J. A., "Application of Automata Theory to Picture Modeling," PhD Dissertation, New Mexico State University, 1976.

11. Whitney, J., "A Pre-Prototype Real-Time Tracking Filter," MS Thesis, New Mexico State University, 1976.

12. Flachs, G. M., W. E. Thompson, R. J. Black, J. M. Taylor, W. Cannon, R. B. Rogers, and Yee Hsun U, "A Pre-Prototype Real-Time Video Tracking System," Final Report for Contract DAAD07-76-C-0024, WSMR, January 1977.

13. Cannon, W., "A Microprocessor System Development Interface," MS Thesis, New Mexico State University, 1977.

14. Rogers, R. B., "Specification for a Two-Pass Symbolic Assembler Generating Absolute Microcode for Arbitrary Microprogram Control Store Architectures," MS Thesis, New Mexico State University, 1977.

15. Perez, P. I. and G. M. Flachs, "A Microprogrammable Processor Architecture for System Design," Technical Report, New Mexico State University, 1977.

16. Ritchell, J., "A General Purpose Microprogrammable Computer Architecture," MS Thesis, New Mexico State University, 1977.

17. Flachs, G. M., W. E. Thompson, J. M. Taylor, W. Cannon, R. B. Rogers, and Yee Hsun U, "An Automatic Video Tracking System," Proceedings of the 1977 National Aerospace and Electronics Conference, May 1977.

GILBERT & *GILES

18. Rogers, R. B. and G. M. Flachs, "Mathematical Modeling
and Simulation in a Programmed Design Methodology," Proceedings of the
First International Conference on Mathematical Modeling, August 1977.

19. Fukunaga, K., A. L. Gilbert, M. K. Giles, O. R. Mitchell,
R. D. Short, and J. M. Taylor, "Segmentation and Structure Analysis for
Real-Time Video Target Tracking," WSMR Technical Report, STEWS-ID-77-1,
October 1977.

20. Flachs, G. M., P. I. Perez, R. B. Rogers, S. J. Szymanski,
J. M. Taylor, and Yee Hsun U, "A Real-Time Video Tracking System,"
Annual Report for Contract DAAD07-77-C-0046, WSMR, January 1978.

21. Gilbert, A. L., "Pattern Recognition and Real-Time Bore-
sight Correction--A Tutorial," SPIE Proceedings, Vol. 134, Photo and
Electro-Optics in Range Instrumentation, March 1978.

22. Giles, M. K. and A. L. Gilbert, "Concepts in Real-Time
Video Tracking," WSMR Technical Report, STEWS-ID-78-1, May 1978.

23. U, Yee Hsun, "Applications of Projection Concepts to
Image Processing," PhD Dissertation, New Mexico State University, 1978.

24. Gilbert, A. L. and M. K. Giles, "A Real-Time Video Track-
ing System," Proceedings of the 24th ISA International Instrumentation
Symposium, May 1978.

15

FIRST BATTLE IN THE HEAT: PHYSIOLOGICAL
LOGISTICS FOR SUCCESS

RALPH F. GOLDMAN, Ph.D.
U.S. ARMY
RESEARCH INSTITUTE OF ENVIRONMENTAL MEDICINE
NATICK, MA 01760

INTRODUCTION:

Even in civilian populations not required to perform hard
work, excess mortality is produced by sudden increases in environmen-
tal heat; when air conditioners were inoperative during a four day
heat wave in the summer of 1970 because of the power outage in New
York City, there was increased mortality (2) despite relatively mild
conditions (T_{db}/T_{wb} = 33°/23°C). The problem is much worse for the
soldier, particularly the infantryman:

> "The life of a foot soldier is divided between two extremes
> of labour and inactivity. Sometimes he is ready to sink
> beneath fatigue, when, having his arms, accoutrements and
> knapsack to carry, he is obliged to make long marches,
> especially in hot or rainy weather..."
>
> John Pringle, 1752
> Surgeon General to the English Army

Experience with such conditions led Medical officers nearly a century
ago to postulate:

> "If possible, choose the cool season for campaigns in warm
> countries."
>
> Andrew Duncan, 1888
> Surgeon to The Bengal Army

The passage of time has not altered these medical comments; if any-
thing, loads are heavier, protective clothing (body armor, chemical

protective clothing) more restrictive, troops must be selected from
a less physically fit population, and they can be moved into hot en-
vironments rapidly, without time for the "seasoning" associated with
slower movement by foot, rail or ship. A personal letter from Viet
Nam in May, 1970, would have offered little new to Surgeons Pringle or
Duncan, other than the use of aircraft:

> "Last week I was up along the Cambodian border. Troop units
> (battalions) are being relocated and reshuffled to new areas;
> often without any or enough canteens. Short, i.e., 1-3 day,
> patrols need continuous air resupply, especially water. Air
> assets are in critical supply and are continuously working.
> In the space of a couple hours I saw 5 men of one squad dust-
> ed off back to the relocated base camp and several others
> from that one company dusted off for heat casualties--all
> within a couple hours on the 6th of May. The temperature
> was hotter than I'd ever experienced...in the adjacent FSB
> someone reported 126°F...as I departed about 4 PM that day,
> I heard there were more heat casualties on the way in.
>
> ...About 1/3 had severe cramps and 2/3 had severe weakness,
> palpitations, fever and near collapse...Incidently, this
> battalion had just (couple days earlier) been moved into
> this area from their old base...They had been fighting in
> relatively easy rice paddies (dry season) and were now re-
> quired to hack their way thru thick jungle...big difference
> in terms of energy required."

<div align="right">

LTC   MC, USA
Viet Nam May 1970

</div>

Given the high work demanded, particularly during the first few days
of engagement with an enemy, and the inability, given U.S. policy of
non-aggression, to choose either the time or place of the first
battle, there is a significant probability that U.S. troop units will
again be required to move rapidly into a hot climate and fight a de-
cisive battle within hours of arrival. The history of heat trauma as
a war experience (15) adequately documents that the problem will not
disappear. Given reasonably fit, trained and equipped troops, there
are two major physiological factors which will seriously degrade the
military operational capacity of the men; optimum performance requires
previous heat acclimatization of the men; it also requires prior plan-
ning to assure delivery--and ingestion--of adequate drinking water.
This paper addresses these two facets of preparation for a units first
battle in the heat, and describes a validated computer prediction
model to assess the effects of any lack of acclimatization and of any
dehydration.

GOLDMAN

HEAT ACCLIMATIZATION:

Heat acclimatization is a term used to describe changes in a number of different physiological parameters; taken all together, these changes which follow exposure to heat produce a major alteration in the soldier's ability to work in the heat (1,9,14). Although the separate elements in the heat acclimatization process can be induced by a variety of techniques, as described below, full heat acclimatization status is most evoked:

  a. with moderate work (e.g., marching with no load at 5.2 km/hr)
  b. with minimum clothing, if not in direct sunlight (to avoid the clothing limitation of evaporative cooling, and thus prolong march time before heat exhaustion occurs)
  c. with "luxus" drinking water (enough to replace sweat losses and thus avoid dehydration)
  d. with an environmental condition (temperature/humidity) at least as severe as that to be encountered operationally.

Acclimatization to less severe environments, or at lower work levels than will actually be encountered operationally, will produce only partial heat acclimatization for the operational situation. However, because it requires about 100 minutes of daily work in the heat to produce full acclimatization (with 50 minutes clearly less beneficial and 200 minutes producing little added benefit), the acclimatization exposures must start with either moderate work or reduced temperature/humidity conditions. This Institute's typical heat acclimatization regimen involves exposure at 49°C (120°F), 20% RH, marching at 5.6 km/hr for 100 minutes (usually two, 50 minute marches separated by a 10' rest break); the men wear shorts and combat boots (or sneakers). They ingest 1/2 canteen of water (500 ml) before the march and then are asked to drink at least 1/4 canteen (250 ml) at the middle and end of each 50 minute march and just before the end of their 10 minute rest. Total water intake is almost 2 full canteens (1.75L) during the 110 minute exposure, which just about equals sweat losses, since sustainable sweat production is about one liter per hour, with short term rates of 3 to 4 L/hr attainable.

In our experience, with probably over 1000 men, even with these optimum clothing and water and modest work conditions, few, if any of our subjects (modestly fit, garrison troops, 18-25 years old) could complete the full 100 minute march on the first day; we limit

3

these experimental exposures when body temperatures exceed 39.5°C
(103°F) or heart rates exceed 180 to 190 b/m and, thereby, experience
no heat stroke problems and only transient heat exhaustion episodes.
Typically, as shown in Figure 1, on the first day our first man is re-
moved after 30-35 minutes (usually unassisted) while the last man is
removed after completing about half the second 50' march. A dramatic
improvement is seen on the second day, with about 1/3 of the men com-
pleting the march. Each day, more men complete the 100 minutes and by
the 6th or 7th day all are considered "fully acclimatized", except the
usual 2 or 3% who appear to be "unacclimatizable"; these are almost
always individuals who have a low maximum work capacity (which may be
a genetic state resistant to physical conditioning, or may, less often,
reflect inadequate physical conditioning). The extended tolerance
times reflect decreases in heart rate, from 175 b/m after 50' of the
first day's walk, to values at the 100th minute on the 4th day of
165 b/m; body temperatures also decrease, from 39.5°C at 50 minutes of
the first day to <39°C at the end of the 6th or 7th day. As a very
crude rule of thumb, about 33% of the total benefit from heat accli-
matization is incurred with one day of exposure, the second day adds
another 15-20%, the third another 10-15% and the fourth through sixth
or seventh contribute about 10%/day to the fully heat acclimatized
status.

Most of the mechanisms contributing to this state are known.
The improved cardiovascular conditioning that is induced by working in
the heat is a substantial contributor; while hard physical condition-
ing in comfortable environments can contribute substantially, and we
estimate that a very fit soldier has the equivalent of about three
days of heat acclimatization (12), the body's ability to deal with the
problem of distributing its available blood supply between working
muscles, where it picks up the heat produced, and a hot, fully vaso-
dilated skin, where it attempts to eliminate this heat, can only be
fully developed by working in the heat. An initial increase in cir-
culating blood volume (hemo-dilution) during heat acclimatization also
contributes to the decreased strain of work in hot environments; with
chronic exposure to work in the heat (e.g., after a month or more)
blood volume appears to return toward baseline levels. Another major
facet of the heat acclimatization process involves conditioning of the
sweat gland, and its control mechanisms, to begin to produce sweat at
a lower skin temperature (i.e., ≤ 35°C) and to produce more sweat per
gland, thus improving the body's potential for sweat evaporative cool-
ing; these effects can be at least partially produced simply by heat,
e.g., in saunas, hot baths or steam rooms. Another facet of the heat
acclimatization process involves an increase in aldosterone levels so

Figure 1. Work tolerance during seven days of heat acclimatization exposure, marching at 5.6 km/h (3.5 mph) for 100 minutes at 49°C (120°F), 20% RH.

that, with the massive sweating, a more dilute sweat is produced which serves to conserve the body's sodium; this mechanism can be at least partially evoked simply by acute restriction of dietary sodium intake, without work or heat.

The benefits of sweat gland conditioning are, of course, a greater benefit in a hot-dry (desert) environment than they would be in a hot-wet one (jungle or the microclimate within a chemical protective clothing system) where evaporative cooling is less limited by sweat availability than by the limited capacity of the humid air to hold additional moisture (16).

DEHYDRATION:

Dehydration produces as dramatic a degradation of work performance in the heat as acclimatization produces an enhancement. As outlined above, water intake in amounts adequate to replace sweat losses can involve supplying 24L/day. As air temperature exceeds $\sim 30°C$ (86°F), even an unclothed man begins to have difficulty eliminating his resting heat production (1 MET = 50 $kcal/m^2 \cdot hr$ or, for an average 70 kg soldier = 90 kcal/h) solely by radiation and convection; as air temperature $\geq 35°C$, all the heat produced by the body must be eliminated by evaporation of sweat. Heat produced by the body during work (15 kcal/min for short periods during an assault) must be either eliminated from the body or stored in it, with body temperature rising $\sim 1°C$ (1.8°F) for every 60 kcal that must be stored. Heat storage of 80 kcal is the usual voluntary tolerance limit, while heat storage levels of 160 kcal, incurred during some of our field operations, resulted in a 50% risk of heat exhaustion collapse; almost no one could continue with body heat storage >240 kcal (when body temperature would be above 41°C). Each gram of sweat evaporated carries away ~ 0.6 kcal of heat, so evaporation of 1L/hr - the maximum sustainable sweat rate - can compensate for 600 kcal/hr of heat production; this is also about the maximum sustainable one hour work rate for an average man. However, if the sweat is not evaporated at the skin surface, the body does not get the full cooling benefit; with reduced permeability uniforms (e.g., the Std B chemical protective clothing system) the body cooling may be only 300 kcal per liter of sweat, produced at the skin, but evaporated at various points between the skin and clothing surface.

Clearly, the demands for drinking water pose a serious logistics problem. Troops can partially compensate for inadequate replacement of sweat losses, up to a point, by limiting their subsequent work, avoiding exposure to sunlight, or in hot crew compartments, and performing essential work during the coolest parts of the day or

night - if the mission and the enemy allow this. However, Adolph (1) and others (10) have clearly documented the effects of dehydration on military performance. Failure to replace 700 ml of sweat loss (i.e., less than one hour's maximum sweat production) produces a 1% dehydration for an average 70 kg (154 lb) soldier. Man may have an extra 1.5L of body water available, since dehydration levels up to 2% seem to produce little problem. However, more severe dehydration clearly limits performance of hard work: a 6% dehydration (a shorfall of only 4.2L), incurred rapidly (i.e., over 6 to 8 hours, so that the body does not have time to partially adapt by intercellular dehydration) results in a non-effective fighting man. For longer exposures (2-3 days) water deficits greater than 10% are intolerable (1,10) and death occurs with 15 to 25% dehydration.

The availability of drinking water will not necessarily insure adequate ingestion. Troops given adequate water supplies have incurred "voluntary dehydration" levels of 5% or more. Thirst apparently is not an adequate stimulus to avoid such "voluntary dehydration", particularly under the pressures of field operations, and there are often problems of palatability because of the water temperature or pretreatment for purity (1). Thus water discipline, which formerly implied command control to limit water intake, today implies that Commanders, by direction, example and supervision at the squad or fire team level, must insure that adequate water is ingested. Small amounts, drunk frequently (e.g., every 20 to 30 minutes), are much better and produce less problems than infrequent large quantities, which may not be well retained. Reports on field trials from Israel, South Africa and the U.K., where heat acclimatized, fit troops were divided into 3 groups, the first given no water, the second with water ad lib and the third required to replace the sweat loss as measured at each halt, all agree that a march which cannot be completed by most men without water, and may be completed only with difficulty when water is given ad lib, can become almost routine when sweat losses are completely replaced.

PREDICTING EFFECTS OF HEAT ACCLIMATIZATION AND DEHYDRATION:

The factors limiting human tolerance for work in the heat are well understood (7). We have developed, and validated (8) a comprehensive model to predict the metabolic heat production (3,13) and the rectal temperature (4) and heart rate (5) responses to work, environment and clothing of fully acclimatized, well hydrated troops. Using the extensive data base on the changes in these responses during acclimatization, that has been built up at Natick during more than

20 years of acclimatizing hundreds of men for chamber studies and
field operations, and additional data from the extensive literature on
heat acclimatization, a subroutine was developed (6) to include the
range of zero to full (i.e., seven days) heat acclimatization in this
prediction model. A literature search and a number of in-house
studies on the effects of varying levels of dehydration were also con-
ducted and these effects are now (11) also able to be incorporated in
our predictions of performance during military operations in hot en-
vironments.

The basic model postulates the existence of a final, equil-
ibrium deep body temperature (T_{re}) at which the body could balance its
heat production [and environmental radiant and convective (H_{R+C}) load
if skin temperature (T_s) is less than the temperature of its surr-
oundings (T_a)] with its total heat losses by evaporation of sweat [and
by radiation and convection if $T_s > T_a$]. The model begins (4) at a
basic T_{re} of 36.75°C; work produces a T_{re} increase of 0.4°C/100 kcal
of metabolic heat production (M); H_{R+C} increases or decreases T_{re} in
direct proportion to the insulation of the uniform (clo) and the diff-
erence ($T_s - T_a$); evaporative cooling occurs as an exponential func-
tion of the difference between the required evaporation (i.e., $M \pm$
H_{R+C}) and the maximum possible evaporation allowed by (a) the evapora-
tive permeability of the uniform worn (i_m/clo) and (b) the difference
between the vapor pressure of sweat at the skin (P_s) and in the air
(P_a). The rectal temperature at which such an equilibrium could be
established may or may not be compatible with life, let alone continu-
ed performance; nevertheless, in both the model and real life, the
rectal temperature will be inexorably driven toward this equilibrium
T_{ref} (4) until heat stroke occurs (at $T_{re} > 41$°C) or until heat ex-
haustion intervenes. The latter is primarily a function of heart rate
(maximum heart rate = 220 b/m minus age, in years), and the model (5)
exploits the link between rectal temperature and heart rate (HR) to
predict heart rate in the heat during rest, work or recovery, with an
experimentally determined, standard error of estimate of six beats per
minute for groups of four or more heat acclimatized, well hydrated men.
Response time delay factors, and rates of rise, for T_{re} and HR during
rest, work and recovery have been developed and included in the model.

The adjustments for heat acclimatization involve a slight
decrease (at most ∿0.5°C) in the initial, resting rectal temperature
(T_{reo}) with heat acclimatization, a loss of 1/2 day of heat acclimati-
zation status for every day without working in the heat, a difference
of about 1.7°C in the T_{ref} between unacclimatized (0 days) and fully
acclimatized (7 days) troops (i.e., 1.5°C plus the 0.2°C decrease in

T_{reo}), but little or no difference in the time constants for rest, work or recovery (6). In summary, unacclimatized troops start with a little higher body temperature ($\sim$0.2°C) and move toward higher T_{ref} levels (by a total 1.7°C) in about the same time as acclimatized troops; they therefore suffer a more rapid rise in T_{re}. There is little or no difference with or without acclimatization in initial heart rate but, despite the difficulty in estimating differences in heart rates which are well beyond the maximum rates than can occur (i.e., >220 b/m), it appears that the greatest difference in heart rate with or without acclimatization is about 40 b/m. The time pattern of heart rate response depends, without reference to acclimatization, on the available cooling power (4) of the environment. In summary, heart rate responses without acclimatization simply reflect the more rapid rise (and thus earlier collapse) associated with the greater strain on the cardiovascular system of attempting to achieve an equilibrium heart rate level of up to 40 b/m above that for a fully acclimatized man. Note that under conditions of extremely high humidity (whether in the natural environment or the microclimate within clothing or crew compartments), the difference between fully and non-acclimatized subjects decreases and reflects primarily the small difference in initial T_{reo}; one should, in fact, expect little or no benefit from the additional sweat production associated with heat acclimatization if the sweat could not be evaporated (16).
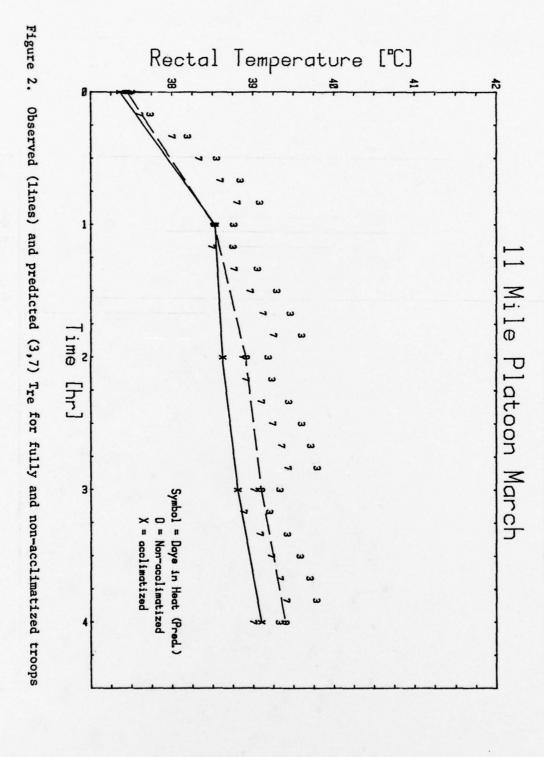
The adjustments for dehydration that have been adopted are somewhat simpler. While these may be changed in the future, subject to additional validating studies, it should be noted that the adjustments for dehydration derived, independently, from several studies did not differ meaningfully, despite the fact that they were conducted with different subjects and levels of dehydration over a two year period. Hydration, over the +3 to -6% range studied, does not appear to alter the equilibrium (i.e., final) T_{re} level very much, if at all. The increase amounts, at most, to one percent, per percent dehydration, of the total difference for the fully hydrated state between the initial (T_{reo}) and equilibrium (T_{ref}) rectal temperature; i.e., if T_{ref}, fully hydrated is 39.35°C [i.e., 36.75°C basic + 2.6°C from M, H_{R+C} and ($E_{req}-E_{max}$) functions], then with 5% dehydration, T_{ref} would only be increased by .13°C (i.e., .05 x 2.6) and this is a meaningless difference. As might be expected, the effects of dehydration on heart rate are more impressive, with the index of heart rate used to compute the actual heart rate increased by 6% per percent dehydration. The time patterns for rates of rise of both rectal temperature and heart rate are also dramatically increased by dehydration, about 10% per percent

dehydration; recovery after work is delayed exponentially, with a 6% dehydration almost doubling recovery time.

APPLICATION OF THE MODEL:

Two field studies, one involving U.S. and the other U.K. troops have been selected for examples of how this predictive model can be used for analysis. The U.S. study (9), a controlled field trial, involved a paratroop platoon, first physically conditioned by walking 8.7 km (14 miles) per day, then all exposed to one day (T1) of a one hour march at 5.6 km/hr at 40°C (105°F) 50% RH. The men then divided into two groups; an "unacclimatized" control group continued to train by marching 8.7 km/day for 12 days (interrupted for weekends) at 18°C (65°F), 50% RH; the other half of the platoon performed the same march at increasing temperatures, the first three days at 40°C, the next at 43.3°, the next at 46° and the last three at 49°C. All men were then exposed again (T2) to the one hour, 5.6 km march at 40°C, 50% RH. The actual difference in rectal temperatures after 45 minutes of T2 was 0.14°C between the heat acclimatized troops and those that simply had been physically conditioned; assuming, as indicated above, that physical conditioning imparts the equivalent of 3 days of heat acclimatization, the predicted difference at the end of 45 minutes is 0.16°C (for the march rate, load carried, uniform worn and environmental exposure used in this chamber phase of the study). The entire platoon was then flown to Fort Kobbe, C.Z. and the next day marched as a unit over an 11 mile trail course at 6.4 km/hr (4 mph), with 10 minute "rest breaks" each hour, during which rectal temperature and heart rate measurements were taken. The model was programmed to predict rectal temperature during the march; the reported T_A (31.1°C/88°F), RH (54%), uniform (fatigues), load (9 kg + clothing), work schedule (50' march, 10' rest) and adequate water (0% dehydration) were used and a dirt terrain (terrain factor 1.1 compared to 1.0 for a blacktop road), 1% grade and 0.89 m/s (2mph) wind were assumed. The predicted rectal temperature responses for troops with 7 days of acclimatization, and with the equivalent of 3 days (induced by physical conditioning), are shown in Figure 2. Despite the confounding of the fact that the actual data was measured during the 10 minute rest, when it is predicted that rectal temperatures should have been dropping rapidly, both the absolute levels of predicted temperatures and the predicted difference between the acclimatized (7 day) and non-acclimatized but physically conditioned ($\simeq$ 3 days) troops show good agreement between the predicted values and those measured 15 years ago.

The data base for the U.K. Study, an operational field

Figure 2. Observed (lines) and predicted (3,7) Tre for fully and non-acclimatized troops

GOLDMAN

trial on the effects of travel fatigue and tropical conditions on the
military efficiency of unacclimatized parachute troops during the
first three days (14), is less adequate for prediction. The troops
were of varied fitness and were judged to be slightly below the aver-
age of the Regiment from which they were drawn. Only task performance,
casualty occurrence and morale were assessed and march rate was not
uniform during the 22.5 km (14 mile) marches which were supposed to be
completed in four hours; the march started within 7 hours of the land-
ing in Singapore after a 43 hour flight with "rest" times of from 7-
1/2 to 14-1/2 hours (average of 11 hours); state of hydration is un-
certain, with ∿15% vomiting; by the end of the second hour, 25% of the
51 men participating were "in difficulty", and 10% had already been
withdrawn; there were more problems with the soldiers feet than with
the heat. Accordingly, while the best estimates for the various in-
put parameters to the model were selected using the actual field op-
eration, the projected results primarily present the opportunity to
contrast the effects of zero and three days of heat acclimatization
with and without progressive dehydration incurred during the march.

A constant march rate (6.4 km or 4 miles per hour) over a
dirt road (1·1 terrain factor), with 30 kg loads at 29.4°C (85°F),
70% RH, were used as inputs to predict rectal temperature (Figure 3)
and heart rate (Figure 4) responses of average troops without accli-
matization and with 3 days of acclimatization (whether induced by pro-
longed physical conditioning or by 3 days of heat acclimatization),
without dehydration, or with progressive dehydration (1st hour = 0,
2nd hour 1%, 3rd hour 2% and 4th hour 3%) either as a result of inade-
quate water availability or inadequate command control to insure that
sufficient water is drunk to replace sweat losses. The clear separa-
tion of the four curves, and reflection on the fact that T_{re} levels
above 39.5°C and HR levels approaching 180 b/m are incompatible with
continuation of activity, provide a strong recommendation for optimum
physical condition, heat acclimatization and hydration of troops ex-
pected to operate effectively during their first battle in the heat.

GOLDMAN

U.K. Army Field March
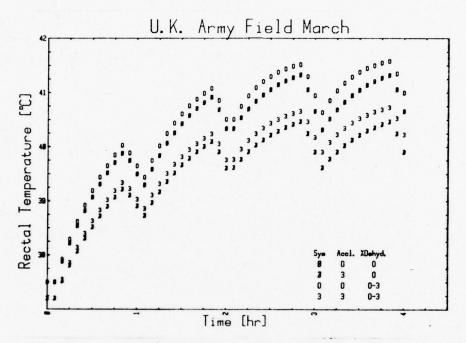
Figure 3. T_{re} w and w/o partial acclimatization and dehydration

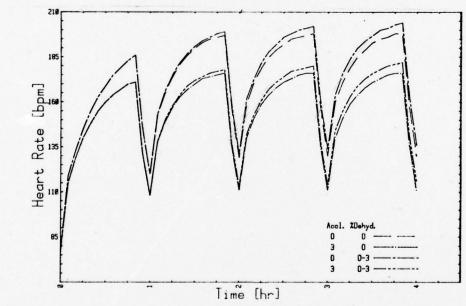
Figure 4. H.R. w and w/o partial acclimatization and dehydration

13

GOLDMAN

1. Adolph, E.F. and Associates. Physiology of Man in the Desert.
Interscience Pub., N.Y., 1947, p 228.

2. Ellis, F.P. Mortality from illness and heat - aggravated illness
in the United States. Environmental Research 5:1-58, 1972.

3. Givoni, B. and R.F. Goldman. Predicting metabolic energy cost.
J. Appl. Physiol. 30:429-433, 1971.

4. Givoni, B. and R.F. Goldman. Predicting rectal temperature res-
ponse to work, environment and clothing. J. Appl. Physiol. 32:812-
822, 1972.

5. Givoni, B. and R.F. Goldman. Predicting heart rate response to
work, environment and clothing. J. Appl. Physiol. 34:201-204, 1973.

6. Givoni, B. and R.F. Goldman. Predicting effects of heat acclima-
tization on heart rate and rectal temperature. J. Appl. Physiol.
35:875-879, 1973.

7. Goldman, R.F. Prediction of human heat tolerance. In: Proc. Int.
Symp. on Adaptations to the Environment, Santa Barbara, CA 1978.

8. Haisman, M.F. and R.F. Goldman. Physiological evaluation of
armored vests in hot-wet and hot-dry climates. Ergonomics 17:1-12,
1974.

9. Joy, R.J.T., R.H. Poe, T.R.A. Davis and E.D. Frohlich. The effect
of prior heat acclimatization on performance in hot climates. Mil.
Med. 129:51-57, 1964.

10. King, J.H. A cavalry detachment three and a half days without
water. Am. J. M. Sci. 75:404, 1878 (cf. Mil. Surg. 74:193-197, 1934).

11. Pandolf, K.B., R.L. Burse, G. Givoni, R.G. Soule and R.F. Goldman.
Effects of dehydration on predicted rectal temperature and heart rate
during work in the heat. Med. Sci. in Sports 9:51, 1977.

12. Pandolf, K.B., R.L. Burse and R.F. Goldman. Role of physical
fitness in heat acclimatisation, decay and reinduction. Ergonomics
20:399-408, 1977.

13. Pandolf, K.B., B. Givoni and R.F. Goldman. Predicting energy expenditure with loads while standing or walking very slowly. J. Appl. Physiol. 43:577-581, 1977.

14. Shepherd, R.D. and J. Barlow. The effects of travel fatigue and tropical conditions on the military efficiency of unacclimatized parachute troops during the first three days. Army Personnel Research Council Report 61/26, July 1962.

15. Whayne, T.F. History of Heat Trauma as a War Exercise. U.S. Army Medical Service Officer Basic Course. Army MSC Graduate School, Army Medical Center, Washington, D.C., 9 June 1951.

16. Wyndham, C.H., N.B. Strydom, J.F. Morrison, F.D. Dutoit and J.G. Kraan. Responses of unacclimatized men under stress of heat and work. J. Appl. Physiol. 6:681-686, 1954.

DIRECT ELECTRONIC FOURIER TRANSFORMS (DEFT) FOR
CAMOUFLAGE SIGNATURE MEASUREMENT (CSM)

Joseph F. Hannigan, Mr.
U. S. ARMY ENGINEER TOPOGRAPHIC LABORATORIES (USAETL)
FORT BELVOIR, VA

INTRODUCTION: This paper presents the theoretical basis and
initial experimental results of a new concept, technique, and method
for measuring and evaluating camouflage in a scientific and quantita-
tive manner. Although methods for precise measurement of color and
contrast have existed for some time, the evaluation of the combined
effect of contrast reduction, patterns, natural foliage, and other
shape disruption methods has remained a subjective judgment on the
part of the camouflage specialist. Thus, the quantitative evaluation
of camouflage on a scientific basis will provide an important and
valuable solution to a unique military problem.

The concept, technique, and method presented herein can be used
to provide the Army with a new capability to scientifically measure
and evaluate this heretofore subjective quantity. It is expected to
provide a valuable tool to the research camouflage specialist in its
present state of development and with further development should pro-
vide essential and valuable information to the military commander in
the field.

The concept and technique is based on the Fourier transform of
an image, and the method uses one of the most recent advances in sur-
face acoustic wave (SAW) technology together with modern communica-
tions technology to obtain the camouflage signature measurement (CSM)
of military objects. This SAW device produces an analog (RF) electri-
cal signal containing the Fourier transform (including amplitude,
frequency, and phase) of the spatial frequencies contained in a
conventional image.

/

This work originated and was performed in-house at the U. S. Army Engineer Topographic Laboratories, under the ILIR program. It is actually a by-product of what has become a long-range program on feature extraction and image analysis. Camouflage is treated as the inverse of feature extraction. It has become difficult to treat one to the exclusion of the other. Hence, reference at time will be made to feature extraction as well as camouflage signature measurement.

CAMOUFLAGE SIGNATURE MEASUREMENT (CSM): A definition of "camouflage" is "the disguising of an installation, vehicle, gun position or ship with paint, garnished nets, or foliage to reduce its visibility or conceal its actual nature or location from the enemy." (1)

The hypothesis taken in this paper is that camouflage is the hiding or disguising of the shape and characteristic features which provide a characteristic signature of a feature or object. In other words, those camouflage measures which cause a change in the spatial frequency pattern of the object are considered in this paper.

The reasoning which led to the concept presented herein is that camouflage and feature extraction are mutual inverses, and that both are signal-to-noise problems. In communications terminology, camouflage is the addition of noise and/or the distortion of the characteristic signature signal of an image while feature extraction is the extraction of a desired signature signal from a noisy signal. Thus, feature extraction becomes a very difficult cross-correlation problem while the measurement of camouflage becomes a relatively less difficult problem of measuring the amount of noise and distortion added to the known signature spectrum, i.e., the overall reduction of the signal-to-noise ratio and the addition or deletion of pronounced spatial frequencies in the signature spectrum.

If a theoretical basis and physical means for measuring the characteristic signature of a feature or object, and the change in the characteristic signature resulting from camouflage can be found and demonstrated; then we will have justified the hypothesis and will have established the basis for a quantitative measurement of camouflage.

THEORETICAL BASIS: The theoretical basis for the concept is the Fourier Theorem and the Fourier transform. Fourier's Theorem has been described as "not only one of the most beautiful results of modern analysis, but also furnishes an indispensable instrument in the treatment of nearly every recondite question in modern physics." (2)

The key Fourier relationships which provide the basic groundwork for camouflage signature measurement (CSM) will be presented. Proofs

of and details on the Fourier transform and its properties can be
found in various texts. A few references are (2) thru (6). Hence,
complete details are omitted in order to take the key formulas and
properties essential to this paper as a point of departure.

The Fourier Theorem states that a function $f(x)$, of spatial per-
iod λ, can be synthesized as a sum of harmonic functions whose wave-
lengths are integral submultiples of λ (i.e., λ, $\lambda/2$, $\lambda/3$. etc.) In
other words, if $f(x)$ is a periodic function of wavelength λ, it can
be represented by a Fourier series of the form:

(1) $f(x) = C_0 + C_1 \cos(2\pi x/\lambda + \phi_1) + C_2 \cos(4\pi x/\lambda + \phi_2) + \ldots$

Another representation of the Fourier series is:

(2) $f(x) = (A_0/2) + \sum\limits_{m=1}^{\infty} A_m \cos mkx + \sum\limits_{m=1}^{\infty} B_m \sin mkx$

where A_m, B_m are the amplitudes of the various harmonic components
as indicated by "m" which has the values $0,1,2,\ldots$, and the wave
number $k = 2\pi/\lambda$ where "λ" is the wavelength or spatial period of the
periodic function, not an optical wavelength.

Another, more compact representation using exponential notation
is:

(3) $f(x) = \sum\limits_{m=-\infty}^{m=+\infty} C_m \exp(jmkx)$

where C_m represents the amplitude of each harmonic component.

The Fourier integral and the Fourier series are alternative ex-
pressions, the integral being the limiting case of the series. It can
be approached from the side of the series by allowing the spatial
repetition to extend to infinity and considering only one period of
the function. Thus, we can write any function, no matter how complex,
as:

(4) $f(x) = (1/2\pi) \int\limits_{-\infty}^{\infty} F(k) \exp(jkx)\, dk$

remembering that we are integrating over submultiples of a single
period, i.e., all k-values become the higher and higher harmonics
required to represent one period of the function.

Now $F(k)$ is referred to as the Fourier transform of $f(x)$ and is

written as:

(5) $\quad F(k) = \int\limits_{-\infty}^{\infty} f(x) \exp(-jkx)\, dx$

The one-dimensional Fourier transform is an important tool and is sufficient for many scientific applications. However, some problems are multidimensional in nature, e.g., image analysis and processing. Therefore, we generalize the Fourier transform to the two-dimensional spatial domain as follows:

(6) $\quad f(x,y,) = [1/(2\pi)^2] \iint\limits_{-\infty}^{\infty} F(k_1, k_2) \exp[j(k_1 x + k_2 y)]\, dk_1\, dk_2$

and

(7) $\quad F(k_1, k_2) = \iint\limits_{-\infty}^{\infty} f(x,y) \exp[-j(k_1 x + k_2 y)]\, dx\, dy$

where k_1 is the wave number in the "x" direction and k_2 is the wave number in the "y" direction.

We now present the Addition Theorem, or superposition property, and the linearity property of Fourier transforms. Any two functions, $af(x)$, and $bg(x)$, will have Fourier transforms $aF(k)$, and $bG(k)$, respectively. Also, the sum of the functions, $af(x) + bg(x)$, has the Fourier transform, $aF(k) + bG(k)$.

(8) $\quad \int\limits_{-\infty}^{\infty} [af(x) + bg(x)] \exp(-jkx)\, dx = \int\limits_{-\infty}^{\infty} af(x) \exp(-jkx)\, dx$

$$+ \int\limits_{-\infty}^{\infty} bg(x) \exp(-jkx)\, dx$$

$$= aF(k) + bG(k)$$

This superposition property means that the spectrum of a linear sum of functions is the linear sum of their spectra. If the spectra are complex, the usual rules of addition of complex quantities apply. Further, any function, as in our case an image, can be regarded as a sum of component parts and the spectrum is the sum of the component spectra.

These separate frequency components, each with their individual specific amplitudes and phases, make up the spatial frequency signa-

ture of the function. Note Eq. (1). Therefore, if the function is
an image, we have the signature of the image. By applying the super-
position principle to the camouflage situation, we take the view that
a military object is one generalized function and camouflage is an-
other generalized function. Hence, the signature spectrum of the
"camouflage," (e.g., patterns, foliage, other shape disruption mea-
sures) is added to the signature spectrum of the object. That is, the
camouflage spatial frequencies, with their individual amplitudes and
phases, are added to the spatial frequencies of the original object.
Note Eq. (8). Actually, some of the original image will probably be
obscured (by foliage, for example). In any case, the result will be
the addition of new frequency components; and, because of phase rela-
tionships and obscuration, the amplitudes of old frequency components
will be increased, decreased, or perhaps eliminated entirely. This
change is a direct result of the addition of the camouflage measure.

   Thus, if an image is changed by the addition of shape disrup-
tive camouflage measures, then the Fourier transform, i.e., the
spectrum of the image, will also be changed accordingly. Hence, if
the frequency signature components can be identified and the change
in amplitude of each signature frequency component can be measured
quantitatively, then the sum of these changes can be used as a quan-
titative measure of the overall effect of the camouflage on the
signature of the object.

   This key point is fundamental to this concept. It is based on
Eqs. (1) thru (8); where any function even one as complicated as an
image, can be specified by a summation of frequencies each having
specific amplitudes and phases. The theoretical basis for the scien-
tific and quantitative measurement of "camouflage" has now been
presented.

METHOD: At this point, we switch from Fourier transform theory to
communications theory and technology. It is now assumed and later
demonstrated that the spatial frequencies of the image can be con-
verted into real RF sinusoidal frequencies as used in standard RF
communications systems. The powers, i.e., the amplitudes and changes
in amplitudes, of RF frequencies for communications applications are
conveniently measured in terms of signal-to-noise ratios (SNR) in
decibels (db). Hence, if we wish to know the strength of a signal
after operating on it in some way, as with an antenna of a certain
size, we simply add the gain of the antenna in db to the strength of
the signal which impinges on the antenna; and conversely, if a receiv-
er amplifier has a certain noise figure (NF), we simply subtract the
NF in db from the strength of the signal. This same approach can now

be adopted in the case of camouflage signature measurement.
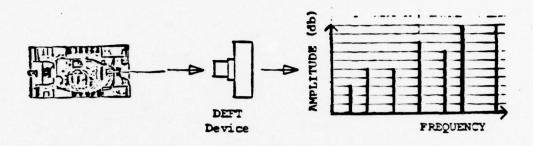
However, since we are dealing with a complete spectrum of frequencies, we should measure the effect on each frequency in the spectrum of the image. This would be logical if the complete image was to be reconstructed. Returning to Fourier theory, we know that the magnitude of any given spatial frequency is an indication of its importance in the image. Therefore, if we select only those frequencies which have a significant amplitude or change in amplitude, we will have reduced the total amount of data, i.e., the spectrum, to be considered, and still have retained the most significant information about the image. Just as we added the separate Fourier frequencies, again recall Eqs. (1) thru (8), to get the original function, we will now add the changes in each principal Fourier frequency component in db to get the "overall" change in the image due to camouflage (i.e., patterns, foliage, etc.). "Overall" is placed in quotation because only the principal frequencies are used.

The method for applying this concept is illustrated schematically in Figure 1. It anticipates the DEFT device and presents the experimental results in an idealized manner based on the theory. In this idealized case, the noncamouflaged tank has seven principal Fourier spectral components, whose amplitudes in db are: 3, 5, 5, 8, 7, 10, and 10. When camouflage is added, the amplitudes of the original spectral components become: 0, 3, 7, 6, 4, 8, and 8 db. In addition, four new components are created as indicated by the dashed lines. These spectral components have amplitudes 3, 5, 6, and 6 db.

The general amplitude reduction of the majority of the original spectral components is attributed to an overall contrast reduction, while the elimination of one spatial component and the addition of four new components is attributed to the addition of new spatial frequencies by the camouflage. The amplitude increase of one component (i.e., the third component went from 5 db to 7 db) is attributed to a reinforcement of that spatial frequency by the camouflage. Since this is a change in the original spatial frequency spectrum, it is considered to be beneficial as far as camouflage, i.e., signature distortion, is concerned. Hence, the total spatial frequency change in the image signature due to camouflage is 36 db. This quantitative measurement could also be expressed as a percentage change due to camouflage, i.e., (36 db) / (48 db) = 75%.

It is anticipated that future work in feature extraction may be able to convert these numbers into a "visibility range" in meters. This would be the distance beyond which the unaided eye would not be

able to "see" the target. The method for a quantitative measurement of camouflage has now been described in an idealized manner.
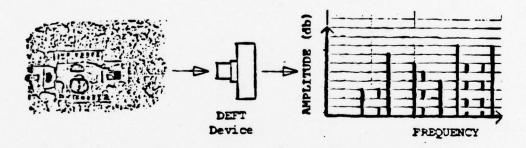
Figure 1. Camouflage Signature Measurement (CSM) Concept. A non-camouflaged tank and a camouflaged tank are viewed by a Direct Electronic Fourier Transform (DEFT) device about the size of a 35-mm camera to obtain the spatial frequency signature of each image. The total change in spatial frequency signature is a quantitative measurement attributable to camouflage.

DEFT TECHNOLOGY: DEFT technology combines acoustics, optics, and electronics in a unique manner to produce the Fourier transform of conventional images directly in the form of an analog (RF) electrical signal. This technology is believed to have a multitude of potential applications in addition to camouflage signature measurement.

The discovery and advancement of DEFT technology is a result of the work of two professors, Dr. P.G. Kornreich and Dr. S.T. Kowel at Syracuse University. Their work was recommended for Army sponsorship by the Advanced Concepts Team (ACT), D.A. The sponsoring agency is

the Night Vision Laboratory with technical support from the US Army
Engineer Topographic Laboratories. A schematic of a device which
represents an early experimental prototype of this technology is
illustrated in Figure 2. This type of device was used for experi-
mental results in this paper.
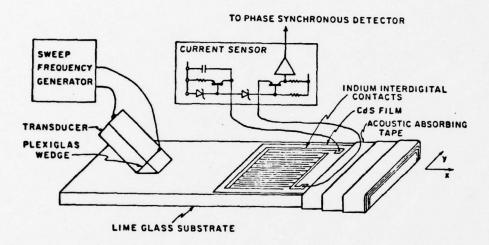
Figure 2. Schematic of an early experimental Direct Electronic
Fourier Transform (DEFT) device. (After Kornreich, P.G., Kowel, S.T.,
et al, "DEFT: Direct Electronic Fourier Transforms of Optical
Images," Proc. IEEE, vol. 62, PP 1072-1087, Aug. 1974).

Significant advances in this technology beyond that illustrated
in Figure 2 have already been made. Devices have been fabricated with
piezoelectric substrates and with dual SAW's for two-directional
scanning at higher frequencies, i.e., 30-MHz range as compared to the
6.0-MHz limit of devices as shown in Figure 2. It can also be recog-
nized that all solid-state construction and low power requirements for
DEFT devices offer great potential for use in the Army field environ-
ment. Present DEFT devices can operate on a 9-volt transistor
battery.

Space does not permit detailed discussion on DEFT technology.
However, technical details can be found in references (7) thru (10).
A very concise description would be to say that the strain induced
by a propagating SAW causes a variation in the conductance of a
photoconductive material. This variation in conductance is spatial
over the whole surface of the photoconductor and also temporal in

accordance with the speed of sound on the surface of the substrate.
When the spatial frequency pattern of the conductance matches that of
one of the component frequencies of the image, then an incremental
analog, a.c. photocurrent at that frequency is generated. These
incremental currents are picked up by a very fine grid of interdigi-
tal electrodes which are deposited with integrated circuit technology.
If the SAW is sinusoidal, swept in frequency, and if synchronous
detection is used, such as a sweep generator together with a network
analyzer, the resulting output is an RF signal with frequencies
corresponding to the spatial frequencies of the image. This output
is shown in references (7) thru (10) to be the Fourier transform of
the image intensity pattern $I_{(x,y)}$. It is obtained in "real time" and
of the form:

$$(9) \quad i_{(ac)} \propto \iint I_{(x,y)} \ \exp[-j(k_1 x + k_2 y)] \ dx \ dy$$

NOTE: Other technologies, such as computer or coherent optical tech-
nologies, are not precluded but have not been investigated for this
application. The Digital Image Analysis Laboratory (DIAL) and the
Recording Optical Spectrum Analyzer (ROSA) could provide candidate
methods for investigation. The size, cost, and complexity of such
systems seem to make them more appropriate for laboratory use than
for general use in the Army field environment.

EXPERIMENTAL RESULTS: The experimental arrangement and close-up
views of the experimental DEFT device are shown in Figure 3. Experi-
ments were performed with scale models (Figure 3.a) and also with
photographs of scale models (Figure 3.b) with varying amounts of
camouflage. The measurement equipment consists of the DEFT device,
the sweep generator, and the network analyzer which can be seen on
the right sides of Figures 3.a and b. A close-up view of the DEFT
device is shown in Figure 3.c, and an internal view is shown in
Figure 3.d.

Sample results obtained with this experimental device are il-
lustrated in Figure 4. This figure shows a tank with increasing
amounts of camouflage from the worst case situation of "zero camou-
flage," i.e., maximum contrast situation of a dark tank against a
white background, to a reduced contrast situation of a dark tank
against a dark background, then a textured background and "camou-
flage foliage" covering the corners of the tank. The "camouflage
foliage" is lichen and the textured background consists of shredded
sponge used by model builders.

The Fourier spatial frequency spectrum (i.e., camouflage signa-

(a)

(b)

(c)

(d)
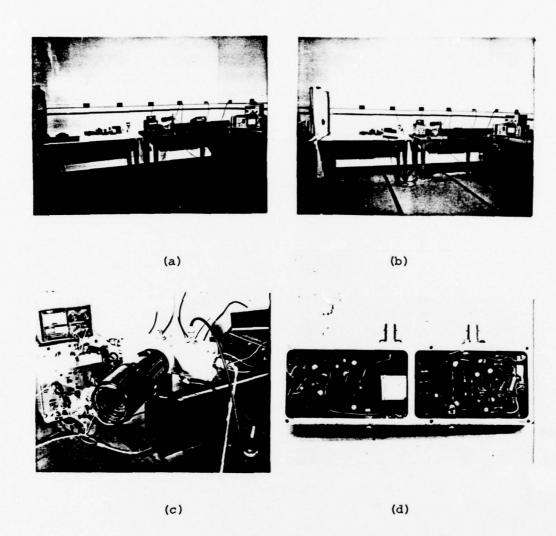
Figure 3. Experimental arrangement and close-up views of the experimental DEFT device used for making Camouflage Signature Measurements in the laboratory with scale models (a) and with photographs of scale models (b). The sweep generator and network analyzer are on the extreme right of (a) and (b). A close-up view of the DEFT device is shown in (c) and an internal view is shown in (d).
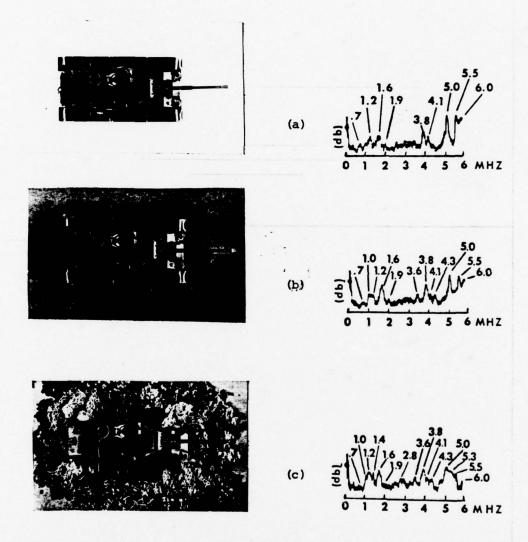
Figure 4. Tanks and corresponding signatures (i.e., Fourier spatial frequency spectra) for Camouflage Signature Measurement: (a) Worst case (zero camouflage) dark tank with white background; (b) Dark tank with dark background; (c) Dark tank with "foliage" background and camouflage foliage on corners.

ture) is shown beside each camouflage configuration. The horizontal axis is frequency from 0.1 to 6.0 MHz which is the bandwidth limit of the DEFT and the vertical axis is amplitude in decibels (db). These camouflage signatures are photographs of the CRT display of the network analyzer. Unfortunately, the graticule is not visible and a great amount of detail is lost in the reduction and reproduction processes. Therefore, the principal frequencies are identified in Figure 4. Only those frequencies whose amplitudes or changes in amplitude are 5 db or more were selected as the principal spatial frequency components.

Table 1 contains a list of the principal spatial frequency components, their amplitudes in terms of SNR based on adjacent noise levels, and the change in SNR of each spatial frequency component. Parentheses are used to indicate certain frequency components that were not considered significant until they became stronger and more significant under the influence of camouflage measures. The initial value, marked with an asterisk, given for these frequencies is based on the initial signal level, i.e., the "noise" level at that frequency, in order to provide a reasonable reference level for subsequent comparison. The need for such a reference is apparent by noting the changes at 5.3 MHz. The plus and minus signs following the change in SNR indicate whether the amplitude of the spatial frequency increased (+) or decreased (-) as a result of "camouflage."

Using the method presented herein, the interpretation to be applied to the numbers contained in Table 1 is as follows: In Case 2 (note Figure 4.b) where there is a contrast reduction due to a better match with the background, there is a 56-db improvement in camouflage (i.e., H-Factor) or a 36% improvement over Case 1 (i.e., 56 db/ 156 db).

In Case 3 (note Figure 4.c) where texture has been added to the background and where the corners together with parts of the front, back, and side edges of the tank have been disrupted with "camouflage foliage," there is a 68-db improvement in camouflage over Case 2 and a 114-db improvement over Case 1. These translate into improvements of 45% (i.e., 68 db/150 db) and 73% (i.e., 114 db/156 db), respectively.

Such types of measurements should be very significant to the tactical field commander and those interested in the "bottom line" for camouflage evaluation. Changes in the individual Fourier spatial frequency components can be expected to provide a great amount of insight for the camouflage specialist in understanding the effect of

Frequency	Case 1 (Fig. 4.a.) S/N Amplitude	Case 2 (Fig. 4.b.) S/N Amplitude (db)	Δ S/N H-Factor Case 1 vs Case 2	S/N Amplitude	Case 3 (Fig. 4.c) Δ S/N H-Factor Case 3 vs Case 2	Δ S/N H-Factor Case 3 vs Case 1
0.7	6	4	2-	0	4-	6-
(1.0)	(5)*	10	5+	14	4+	9+
1.2	10	8	2-	16	8+	6+
(1.4)	(0)*	(0)*	(0)*	8	8+	8+
1.6	10	16	6+	12	4-	2+
1.9	6	4	2-	0	4-	6-
(2.8)	(1)*	(1)*	(0)*	7	6+	6+
(3.6)	(2)*	7	5+	7	0	7+
3.8	13	14	1+	14	0	1+
4.1	8	2	6-	2	0	6-
(4.3)	(2)*	6	4+	6	0	4+
5.0	30	24	6-	20	4-	10-
(5.3)	(6)*	10	4+	16	6+	10+
5.5	29	23	6-	14	9-	15-
6.0	28	21	7-	10	11-	18-
	156 db	150 db	56 db (36%)		68 db (45%)	114 db (73%)

Table 1. Principal spatial frequency components, their amplitudes (S/N) in decibels, changes in amplitudes and overall camouflage evaluation, i.e., H-Factor, at bottom of Case 2 and Case 3 columns.

13

various camouflage measures. It is also by studying the structure
and changes in structure of these same spatial frequencies that in-
sights into feature extraction can be expected.

CONCLUSIONS: It is concluded that a valid theoretical basis and a
valid method have been presented for the scientific and quantitative
measurement of those camouflage measures which are designed to disrupt
the shape of an object or reduce its contrast with its background.
It is further concluded that DEFT technology makes the method adapt-
able for use in the tactical Army environment as well as the research
laboratory. It is believed that this concept and method opens the
door and is an initial step toward making camouflage a military
science as well as a military art. As such, this effort is believed
to be a significant advance toward fulfilling the words of Lord
Kelvin, "When you can measure what you are speaking about, and express
it in numbers, you know something about it; but when you cannot
express it in numbers, your knowledge is of a meager and unsatisfac-
tory kind ...."

A word of caution is required. This concept, technique, and
method requires understanding on the part of the user. It cannot be
applied blindly, without training and experience in camouflage. It
is actually a sophisticated tool, and like all tools, it requires
understanding and judgment in its use. In the hands of the skilled
camouflage specialist, it could lead to new insights into what pro-
duces "good" camouflage and perhaps a scientific basis for the design
of new camouflage measures.

HANNIGAN

(deceased) and G.H. Sigel (retired), former Section Chiefs, Camouflage Laboratory, for an awareness of the need for a scientific and quantitative measurement of camouflage and to all other associates whose discussions have enhanced the preparation of this paper.

REFERENCES

1. Webster, D., "Webster's Third New International Dictionary of the English Language Unabridged," G. & C. Merriam Co.

2. Campbell, G.A., Foster, R.M.,"Fourier Integrals for Practical Applications," D. Van Nostrand Co., Inc. p. 3.

3. Bracewell, Ron, "The Fourier Transform and Its Applications," McGraw-Hill Book Co.

4. Panter, P.F., "Modulation, Noise, and Spectral Analysis," McGraw-Hill Book Co., pp 8-62.

5. Andrews, H.C., "Computer Techniques in Image Processing," Academic Press.

6. Goodman, J.W., "Introduction To Fourier Optics," McGraw-Hill.

7. Kornreich, P.G., Kowel, S.T., et al, "DEFT: Direct Electronic Fourier Transforms of Optical Images," Proc. IEEE, Vol. 62, pp 1072-1087, Aug. 1974.

8. Kowel, S.T., Kornreich, P.G., et al, "Two-Dimensional Direct Electronic Fourier Transform (DEFT) Devices, Analysis, Fabrication, and Evaluation," Report DAAG 53-76-C-0162, 30 June, 1977, Syracuse University, Army Night Vision Laboratory, Fort Belvoir, VA.

9. Kowel, S.T., Kornreich, P.G., et al, "Two-Dimensional Fourier Imaging of Light Using Acoustic Psuedo Beam Steering," 1975 Ultrasonic Symposium Proceedings, IEEE Cat. #75 CHO 994-4SU pp. 136-140.

10. Kowel, S.T., Kornreich, P.G., et al, "Experimental Confirmation of Two-Dimensional Acoustic Processing of Images," 1976 Ultrasonic Symposium Proceedings, IEEE Cat. #76 CHII 20-5SU pp. 668-672.

THE STRANGE BEHAVIOR OF ELECTROMAGNETIC WAVES IN CONDUCTING MAGNETO-DIELECTRIC MEDIA

ORVILLE R. HARRIS, PhD
PROFESSOR EMERITUS, UNIV. OF VA.
U. S. ARMY FOREIGN SCIENCE AND TECHNOLOGY CENTER
CHARLOTTESVILLE, VA 22901

1. Introduction

Reflection and transmission of electromagnetic waves at boundaries between dielectric media, and reflection at highly conducting boundaries have been treated extensively in the literature for years. Usually approximations have been made to fit the special case under discussion. When dealing with good dielectrics, the conductivity is usually neglected and for highly conducting materials the dielectric constant is neglected. Almost always, the permeability is considered to be equal to that of free space. Between these limiting cases, there are a host of materials in which all three constitutive parameters, permeability, permittivity, and conductivity should be considered.

The fact that any one or all three of these parameters can be complex generates the following interesting situation. Assume $\varepsilon = \varepsilon' - j\varepsilon''$, $\sigma = \sigma' - j\sigma''$ and $\mu = \mu' - j\mu''$. In a conducting medium the apparent dielectric constant is $\varepsilon_T = \varepsilon - j\sigma/\omega = (\varepsilon' - \sigma''/\omega) - j(\varepsilon'' + \sigma'/\omega)$, (1) where $\omega = 2\pi f$. Thus, if σ' and σ'' are of the same sign as ε' and ε'', the real part of dielectric constant appears to be decreased. Further, Landau and Lifshitz[1] have shown that μ'' and ε'' must always be positive, but that there is no physical restriction on the sign of μ' or ε'. It can also be shown that σ' must also always be positive by noting that the total heat dissipated per unit time and volume, in a conducting medium is

$$Q = \omega/2 \; \left| \mu'' \tilde{H}^2 + \varepsilon'' \tilde{E}^2 + \frac{\sigma'}{\omega} \tilde{E}^2 \right| \tag{2}$$

where $\sim$ indicates the time average of the fields. Hence, σ', μ'', and ε'' must all be positive to insure that the medium heats up rather than cools. In this paper the following constitutive parameters will be used.

$$\mu = \mu_o(\mu_r' - j\mu_r'') = \mu_o|\mu_r|e^{j\Theta\mu} \tag{3}$$

$$\varepsilon_T = \varepsilon_o(\varepsilon_r - j\sigma/\omega\varepsilon_o) = \varepsilon_o|\varepsilon_r|e^{j\Theta\varepsilon} \tag{4}$$

$$= \varepsilon_o\varepsilon_r(1 - j\tan\delta) \tag{5}$$

Here $\mu_o = 4\pi \times 10^{-7}$ H/m, and $\varepsilon_o = (1/36\pi) \times 10^{-9}$ F/m in MKSQ dimensions and units. The term $\tan\delta$ is the loss tangent. The subscript r denotes the dimensionless relative value. In many dielectric materials ε_r and $\tan\delta$ are well documented, but little is known about μ'' in magnetic materials. The term ε_r, where necessary, will be taken as $(\varepsilon' - \sigma''/\omega\varepsilon_o)$ which may mean that ε_r can appear to be negative in some materials. This occurs in some materials such as gold at infrared wave lengths, as will be shown later. When ε_r is negative it is usually not possible to determine whether ε' or $\sigma''\omega\varepsilon_o$ is the predominant factor.

Equations A(8) and A(9) in the appendix define the propagation constant in conducting media. Setting $\mu_{rn} = 1 + j0$ and $\sigma_n = 0$ then

$$k = \omega^2\mu_o\varepsilon_o = \omega^2/c^2 = \beta \text{ rad/m} \tag{6}$$

where a wavelength is defined as

$$\lambda = 2\pi/\beta = 2\pi/k \tag{7}$$

in which all quantities are real. If the medium is conducting and k becomes complex, do we define λ as $2\pi/\beta$, or as $2\pi/k$? In this paper the latter definition is used which dictates that the wavelength (λ), index of refraction (n), and phase velocity of propagation (υ_p) must all be complex quantities. Also, since λ is complex, all distances must be complex so that the total phase change along any path length can be a real number of radians.

2. Weakly Conducting Dielectrics at Extremely Low Frequencies (ELF)

Although the title of this paper refers specifically to conducting media it is necessary to begin with two semi-infinite non-conducting dielectrics in contact in order to describe several interesting angles of incidence as defined by Lytle and Lager[2]. To begin with, a vertically polarized plane wave incident on the boundary (Figure 5) between free space and a perfect dielectric is assumed. When the medium of incidence (medium 1) is free space, there is an angle of incidence (Θ_i) such that total transmission into a medium of refraction (medium 2) occurs, and the reflection coefficient $R_{12}^H = 0$(eqn.A(23)). This is the well known Brewster angle Θ_B which occurs when:

$$\Theta_B = \Theta_{t2} = \sin^{-1} \left| \varepsilon_{r2}/(\varepsilon_{r2}+1) \right|^{-\frac{1}{2}}; \quad \varepsilon_{r1} > \varepsilon_{r1}, \varepsilon_{r1} = 1. \qquad (8)$$

When the situation is reversed such that the medium of incidence (medium 1) has a large dielectric constant and the medium of refraction is free space (medium 2), three more interesting angles of incidence occur as pointed out by Lytle and Lager:

$$\Theta_B' = \sin^{-1} \left| 1/(\varepsilon_{r1}+1) \right|^{\frac{1}{2}}; \quad \varepsilon_{r2} = 1 \qquad (9)$$

at which total transmission into free space occurs ($R_{12}^H = 0$) and the angle of transmission (Θ_{t2}) is equal to the Brewster angle (Θ_B) above,

$$\Theta_c = \sin^{-1} (1/\varepsilon_{r1})^{\frac{1}{2}}; \quad \varepsilon_{r2} = 1 \qquad (10)$$

called the critical angle, which occurs when the reflection coefficients for both vertical and horizontal polarization are both equal to unity (i.e. $R_{12}^H = R_{12}^E = 1$), and

$$\Theta_D = \sin^{-1} \left| (\varepsilon_{r1}+1)/(\varepsilon_{r1}^2+1) \right|^{\frac{1}{2}}; \quad \varepsilon_{r2} = 1 \qquad (11)$$

called the "Devil" angle, which occurs when the reflection coefficient is purely imaginary (i.e. $R_{12}^H = -j$). The name was coined by Lytle and Lager because of their devilment at the unusual occurrences at this angle.

This latter angle is treated, as example 1, in detail in Figure 1 where the medium of incidence has constitutive parameters of $\varepsilon_{r1} = 25$, and $\sigma_1 = 0$. The frequency (f) used in this example is one Hertz since this represents a frequency near the low end of the spectrum. However, the results here and in subsequent examples will be obtained at any frequency provided the loss tangent (tan $\delta = \sigma/\omega\varepsilon_T$) Eqn. A(11) remains constant as a function of frequency. The H-mode transmission and reflection coefficients (T_{12}^H and R_{12}^H) are calculated using equations A(22) through A(25) in the appendix where the "Devil" angle of evidence is 11.759°. The E-mode coefficients are not calculated since they are uniquely determined by equations A(26) through A(30). The principle points of interest to be observed are as follows:

(a) The reflected H-field vector is equal in magnitude to the incident field but lags the latter in time phase by 90°.

(b) The transmitted H-field is 1.414 times greater than the incident field and lags the latter in time phase by 45°.

(c) The angle of refraction (Θ_{t2}) is a complex angle of 90°+j 11.07° which is quite contrary to what would be expected.

(d) The planes of constant phase and constant amplitude in free space (medium 2) are not coincident. The constant phase plane is

perpendicular to the interface, while the constant amplitude is plane parallel to the interface.

(e) The incident angles Θ_B', Θ_C, and Θ_D all occur within the range of $\Theta_{i1} = 11^\circ$-12° such that $\Theta_B < \Theta_C < \Theta_D$ (Figure 3). For all angles below the critical angle (Θ_c) the transmitted H-field (H_{t2}) is in phase with the incident H-field (H_{i2}). Between Θ_B' and Θ_C the reflected H-field (H_{r1}) is out of time phase with the incident field by 180°.

(Note 1 – When the angle of refraction is complex, the planes of constant phase and amplitude, in the medium of refraction (medium 2), are not usually coincident. We assume in all examples that they are coincident in the medium of incidence (medium 1). The angle (ψ) of constant phase plane can be calculated from

$$\cos \psi = \frac{\beta_2 Rcos + \alpha_2 Icos}{\left| (\beta_2 Rcos + \alpha_2 Icos)^2 + (\beta_2 Rsin + \alpha_2 Isin)^2 \right|^{\frac{1}{2}}} , \qquad (9)$$

where $\cos \Theta_{t2} = Rcos + jIcos$, and $\sin \Theta_{t2} = Rsin + jIsin$ are the complex cosine and sine of the complex angle of refraction (Θ_{t2}). The phase constant (β_2), and attenuation constant (α_2) are defined by equation A(9). Similarly, the angle ($\emptyset$) of the constant amplitude plane in medium 2 can be determined by
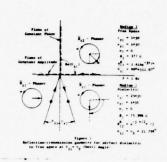
$$\cos \emptyset = \frac{\beta_2 Icos - \alpha_2 Rcos}{\left| (\beta_2 Icos - \alpha_2 Rcos)^2 + (\beta_2 Isin - \alpha_2 Rsin)^2 \right|^{\frac{1}{2}}} , \qquad (10)$$
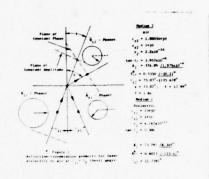
Both of these angles are measured from the plane of incidence (Figure 5) and can be considered as the refraction angles for constant phase and constant amplitude propagation directions. The respective planes are normal to these directions.)
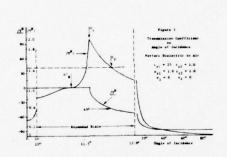
When the two media are allowed to become weakly conducting, as shown in Figure 2, the situation changes quite drastically. The principal points to note are:
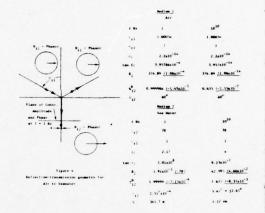
(a) The angle of refraction (Θ_{t2}) has changed, but is still complex. The real part of Θ_{t2} still corresponds to the refraction angle (ψ) of constant phase planes.

(b) The reflected H-field has decreased in magnitude and now lags the incident field by 153.4° in time phase. The transmitted H-field has also decreased in magnitude but lags in time phase by only 30.2°. It should be noted that if the conductivity of medium 1 is increased (say $\sigma = 10^{-6}$ S/m), the reflected field amplitude becomes unity and lags by 180°. Hence the transmitted field becomes zero in magnitude so that the boundary behaves like a near perfect magnetic sheet reflector.

Figure 1
Reflection-transmission geometry for perfect dielectric to free space at $\theta_{i1} = \theta_B$ (Brewi Angle)

Figure 2
Reflection-transmission geometry for lossy dielectric to air at $\theta_{i1} = \theta_B$ (Brewi Angle)

Figure 3
Transmission Coefficient vs Angle of Incidence
Perfect Dielectric to Air
$\epsilon_{r1} = 25$ $\epsilon_{r2} = 1.0$
$\mu_{r1} = 1.0$ $\mu_{r2} = 1.0$
$\sigma_1 = 0$ $\sigma_2 = 0$

Figure 4
Reflection-transmission geometry for Air to Seawater

3. Strongly Conducting Media

Sea water is a well known example of a strongly conducting medium. Much interest has been shown in the reflection-transmission coefficients when an electromagnetic wave is incident on the surface of an ocean (medium 2) from air (medium 1). The constitutive parameters for both media at one Hertz and 10 GHz were estimated from an extensive literature search and are shown in figure 4.

It comes as no surprise that sea water is a good reflector of electromagnetic energy. However, it is not well known that the refracted H-field (H_{t2}) is nearly twice that of the incident field for angles of incidence up to 60°. The reason for this is that the incident, reflected, and refracted H-fields are all in phase. This does not mean that a lot of energy is refracted, however, since the refracted E-field amplitude is reduced by a factor of $|Z_2/Z_1|$ which is

5

of the order of 5×10^{-6} at one Hertz. This suggests that, if one de-
sired to communicate with a submerged submarine at one Hertz, an
H-field sensor, such as a magnetometer, may be more effective than an
E-field sensor such as a trailing wire antenna.

It is also interesting to note that the angle of refraction
is nearly 90° for both the constant phase and constant amplitude pro-
pagation at most angles of incidence at low frequencies. This is true
at all frequencies for the constant amplitude planes, however at 10
GHz the constant phase plane angle of refraction is 5.6°. Thus, the
energy travels straight down in the sea water.

There is one exception to the fact that the reflection co-
efficient (R_{12}^{H}) is nearly unity at all angles of incidence. This
exception is at an angle of incidence of 88°-89°, depending on fre-
quency, where a quasi-Brewster angle occurs such that the reflection
coefficient becomes quite small and the transmission is nearly unity.
This effect is more pronounced at the higher frequencies.

Figure 4 also shows that the skin depth(s), where $H/H_0 = e^{-1}$,
is 30 000 times greater at one Hertz than at 10 GHz. At fifty Hertz
the skin depth(s) is 48.3 meters as compared to 341.7 m at one Hertz.
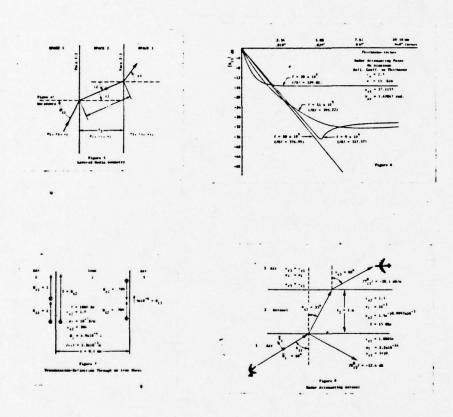
4. Layered Conducting and Permeable Media

The subject of multi-layered sandwiches is of extreme inter-
est in applications such as radar radomes, radar attenuating paints,
radar attenuating aerosols, and electromagnetic wave shielding. Born
and Wolf[3]; Cady, Karelitz, and Turner[4]; and others have treated this
subject in great detail. Born and Wolf, on the one hand, were primar-
ily interested in optics so several approximations were made. On the
other hand, Cady, et al. were interested in microwaves through good
dielectrics, so somewhat different approximations were made. Both
approaches to the problem involved chain matrix multiplication of a
succession of 2x2 matrices. If one follows this approach but makes no
approximations, a matrix equation such as equation A(36) results.
Using this, the reflection coefficient at the incident face of a multi-
layered sandwich and the transmission coefficient through the last
face can be calculated by virtue of equations A(37) through A(47) at
any angle of incidence. These equations are also applicable for any
value of ε_r, μ_r, and σ providing ε_r and σ are real. The permeability
can be complex as indicated by equations A(7) and A(8). The geometry
of a single layer sandwich is shown in figure 5. It is obvious how to
add additional layers up to $m = n-1$ where n is the number of faces.

Consider as the reflection coefficient of a radar attenuating
paint on aluminum for normal incidence at frequencies of 9 GHz, 10 GHz,
11 GHz, and 20 GHz. The constitutive parameters of the paint, used in
the problem, are shown in figure 6. It is not known whether such a
paint having these characteristics can be compounded, but, at least,

6

HARRIS

here is one set of parameters which will produce the results shown.
They were selected such that $|(\mu_r/(\epsilon_r - j\sigma/\omega_{to}))|^{\frac{1}{2}} = 1$, so that the
intrinsic impedance at normal incidence and $f = 10^{10}$ Hz matches that
of air. The reflection coefficients, at normal incidence, are plot-
ted as a function of paint thickness. Note that at a thickness of
5.08 mm (0.02"), the magnitude of the reflection coefficient is 40 dB
below that of a perfect reflector at a frequency of 10 GHz.

Now consider a thin sheet of iron, as an electromagnetic
shield at one KHz, as shown in figure 7. The most salient points to
note are as follows:

(a) The refracted H-field, at the incident face, is twice
the incident field as occurs in all highly conducting fields for the
same reasons stated earlier.

(b) The H-field exiting through the output face is negligi-

7

ble because the reflected field lags incident field by 180°. They
therefore cancel each other at the output interface.

As a final example consider a one meter thick aerosol be-
tween a radar, operating at 15 GHz, and an aircraft flying overhead.
The constitutive properties of this aerosol are shown in figure 8.
It is not known whether such an aerosol could be compounded but the
principle is demonstrated that, with such an aerosol at an angle of
incidence of 60°, the overall transmission coefficient magnitude is
-28.1 dB per meter thickness. A signal reflected from the aircraft,
upon returning to the radar would be subject to the same degree of
attenuation which could render an aircraft nearly undetectable.

5. Highly Conducting Media

In addition to the basic metals there are many highly con-
ducting materials such as graphite loaded epoxy and conducting poly-
mers. Little is known about their constitutive parameters so that
the equations included herein cannot be used to investigate various
practical examples. Examination of the published data on many materi-
als, however, leads one to suspect that conductivities become signifi-
cantly complex at far infrared frequencies. The complex index of re-
fraction of the basic metals are well documented so that this suspi-
cion can be verified and at least demonstrate that complex conductivi-
ties do indeed exist.

We consider the case of gold which has a complex index of
refraction of $n = n'-jn'' = 25.2 -j55.9$ at $\lambda = 9.9$ μm. Noting that
$n = c/\omega$ $(\beta-j\alpha)$, the propagation constant $(k = \beta-j\alpha)$ can be calculated
as $k = 2.52 \times 10^7 -j3.41 \times 10^7$ at a frequency of 3.03×10^{13} Hertz. Find-
ing the square of k, where $k^2 = \omega^2\mu\epsilon-j\omega\mu\sigma$, allows us to calculate σ if
we assume values of unity for μ_r and ϵ_r. Thus

$$k^2 = -5.31 \times 10^{14} -j1.718 \times 10^{14} = \frac{\omega^2}{c^2} \mu_r \epsilon_r \ (\frac{-\sigma''}{\omega\epsilon_o} -j \frac{\sigma'}{\omega\epsilon_o})$$

from which σ is determined as

$$\sigma = \sigma'-j\sigma'' = 4.274 \times 10^6 -j1.605 \times 10^6$$

$$= 4.565 \times 10^6 \ /-0.359 \ rad \ S/m$$

which is an order of magnitude smaller than the direct current value
of $(\sigma \approx 4 \times 10^7)$S/m for gold.

6. Conclusions

Unfortunately, in a paper of this length, it is impossible
to include derivations of the equations used to make the calculations

of the quantities shown in the various figures. In fact, it is not possible to show all of the quantities calculated on each figure. However, the most significant quantities, which explain the strange behaviors discussed, are included. Also all equations used are included.

This by no means exhausts the subject of strange behaviors under other circumstances. However, the author has found that when all parameters are included, no matter how small, many surprises can occur under the most ordinary of circumstances.

Acknowledgements

References

1. Landau, L. D., E. M. Lifshitz, "Electrodynamics of Continuous Media," Pergamon Press, Inc., 1975 (Reprint) pp 251-252, N.Y.

2. Lytle, R. J., D. L. Lager, "Electromagnetic Propagation, Transmission, Reflection, and Refraction: Equations and Numerical Results." Report No. UCRL-51245, Lawrence Livermore Laboratory, University of California, Livermore, California 94550, 20 June 1972.

3. Born, Max, Emil Wolf, "Principles of Optics," Pergamon Press, 5th Ed., 1975, N.Y.

4. Cady, W. M., M. B. Karelitz, L. A. Turner, "Radar Scanners and Radomes," MIT Radiation Laboratory Series, Vol 26, McGraw-Hill 1948, N.Y.

Appendix
Mathematical Equations

Snell's Laws

$$\cos \Theta_{i1} = \cos \Theta_{r1} \qquad \text{A (1)}$$

$$k_1 \sin \Theta_{i1} = k_2 \sin \Theta_{t2} \qquad \text{A (2)}$$

Fresnel's Equations

T. E. mode (horizontal polarization)

$$\overline{E}_t = \frac{2\mu_2 k_1 \cos \Theta_i}{\mu_2 k_1 \cos \Theta_i + \mu_1 (k_2^2 - k_1^2 \sin \Theta_i)^{1/2}} \overline{E}_i \qquad \text{A (3)}$$

$$\overline{E}_r = \frac{\mu_2 k_1 \cos \Theta_i - \mu_1 (k_2^2 - k_1^2 \sin^2 \Theta_i)^{1/2}}{\mu_2 k_1 \cos \Theta_i + \mu_1 (k_2^2 - k_1^2 \sin^2 \Theta_i)^{1/2}} \overline{E}_i \qquad \text{A (4)}$$

T. M. mode (vertical polarization)

$$\overline{H}_t = \frac{2\mu_1 k_2^2 \cos \Theta_i}{\mu_1 k_2^2 \cos \Theta_i + \mu_2 k_1 (k_2^2 - k_1^2 \sin^2 \Theta_i)^{1/2}} \overline{H}_i \qquad \text{A (5)}$$

$$\overline{H}_r = \frac{\mu_1 k_2^2 \cos \Theta_i - \mu_2 k_1 (k_2^2 - k_1^2 \sin^2 \Theta_i)^{1/2}}{\mu_1 k_2^2 \cos \Theta_i + \mu_2 k_1 (k_2^2 - k_1^2 \sin^2 \Theta_i)^{1/2}} \overline{H}_i \qquad \text{A (6)}$$

The subscripts i, r, and t, denote incidence, reflection, and transmission respectively. The numerical subscripts 1, 2, ---, n identify the medium in which the wave is traveling. Where the subscript r is used with the permitivity (ϵ), and permeability (μ) it denotes the relative value such that $\epsilon_n = \epsilon_{rn} \epsilon_o$, and $\mu_n = \mu_{rn} \mu_o$ where

$\epsilon_o = \frac{1}{36\pi} \times 10^{-9}$ F/m, and $\mu_o = 4\pi \times 10^{-7}$ H/m in rationalized MKS units.

The relative values are dimensionless. The conductivity (σ_n) has the dimensions of Siemens/meter (S/m).

Horizontal polarization (T. E. mode) equations (3), and (4) apply when the incident $\overline{E}$-field is parallel to the boundary, as shown in figure 1. For vertical polarization (T. M. mode), equations (5), and (6) apply when the incident $\overline{H}$-field is parallel to the boundary. In these equations, the relative permeabilities are allowed to be complex so that

$$\mu_n = \mu_{rn}\mu_o = |\mu_{rn}| e^{j\Theta_u} \mu_o \tag{A (7)}$$

The square of the propagation constants which result from the wave equations are

$$k_n^2 = \omega^2 \mu_n \epsilon_n - j\omega\mu_n\sigma_n$$
$$= \omega^2 \mu_o \epsilon_o \left[\mu_{rn}\left(\epsilon_{rn}' - j\frac{\sigma_n}{\omega\epsilon_o}\right)\right] \tag{A (8)}$$

where $\omega = 2\pi f$, and f is the frequency in Hertz. The propagation constant is expressible as

$$k_n = \beta_n - j\alpha_n \tag{A (9)}$$

where β_n is the phase constant in radians per meter, and α_n is the attenuation constant in nepers per meter. The quantity ϵ_{rn} can be complex as well, such that

$$\epsilon_{rn} = \epsilon_{rn}' - j\left(\epsilon_{rn}'' + \frac{\sigma_n}{\omega\epsilon_o}\right)$$
$$= |\epsilon_{rn}| e^{j \tan \delta} \tag{A (10)}$$

In this report the quantities in parentheses are treated as one quantity attributable to a single equivalent conductivity. The loss tangent then is

$$\tan \delta = \frac{\sigma_n}{\omega\epsilon_T} \tag{A (11)}$$

where

$$\epsilon_T = \epsilon_o |\epsilon_{rn}| \left(1 - j\frac{\sigma_n}{\omega|\epsilon_{rn}|\epsilon_o}\right) \tag{A (12)}$$

If ϵ_{rn} and $\tan \delta$ are known at a given frequency, the conductivity can be calculated by

$$\sigma_n = \frac{1}{18} f_m |\epsilon_{rn}| \tan \delta \times 10^{-9} \tag{A (13)}$$

where f_m is the reported frequency at which $\tan \delta$ was measured. Here, σ_n is calculated from the loss tangent at the nearest reported frequency to the actual frequency of interest and then treated as a constant value.

HARRIS

From equations (3) and (4) the TE mode transmission (T) and reflection (R) coefficients for a positive going wave can be expressed as

$$T_{12}^E = \frac{\overline{E}_{t2}}{\overline{E}_{i1}} = \frac{2}{1 + Y_{12}^E} \qquad\qquad A\ (14)$$

and

$$R_{12}^E = \frac{\overline{E}_{r1}}{\overline{E}_{i1}} = \frac{1 - Y_{12}^E}{1 + Y_{12}^E} \qquad\qquad A\ (15)$$

so that

$$R_{12}^E = T_{12}^E - 1 \qquad\qquad A\ (16)$$

where the admittance is

$$Y_{12}^E = \frac{1}{Z_{12}^E} = \frac{\mu_1 k_2}{\mu_2 k_1} \frac{\left[1 - (k_1/k_2 \sin \Theta_{i1})^2\right]^{1/2}}{\cos \Theta_{i1}} \qquad\qquad A\ (17)$$

and Θ_{i1} is the angle of incidence from medium 1 on face 1-2. These equations hold at any interface between two media for a positive going incident wave which is from left to right in figure 1. For a negative going incident wave, these equations become

$$T_{21}^E = \frac{\overline{E}_{t1}}{\overline{E}_{i2}} = \frac{2}{1 + Y_{21}^E} \qquad\qquad A\ (18)$$

and

$$R_{21}^E = \frac{\overline{E}_{r2}}{\overline{E}_{i2}} = \frac{1 - Y_{21}^E}{1 + Y_{21}^E} \qquad\qquad A\ (19)$$

so that

$$R_{21}^E = T_{21}^E - 1 \qquad\qquad A\ (20)$$

where

$$Y_{21}^E = \frac{1}{Z_{21}^E} = \frac{\mu_2 k_1}{\mu_1 k_2} \frac{\left[1 - (k_2/k_1 \sin \Theta_{i2})^2\right]^{1/2}}{\cos \Theta_{i2}} \qquad\qquad A\ (21)$$

At successive boundaries, the subscripts change from 12 and 21 to 23 and 32 at face 23 as shown in figure 1.

HARRIS

For the T. M. mode, similar equations result as follows.

$$T^H_{12} = \frac{\overline{H}_{t2}}{\overline{H}_{i1}} = \frac{2}{1+Y^H_{12}} \qquad \text{A (22)}$$

$$R^H_{12} = \frac{\overline{H}_{r1}}{\overline{H}_{i1}} = \frac{1-Y^H_{12}}{1+Y^H_{12}} \qquad \text{A (23)}$$

$$R^H_{12} = T^H_{12} - 1 \qquad \text{A (24)}$$

$$Y^H_{12} = \frac{1}{Z^H_{12}} = \frac{\mu_2 k_1}{\mu_1 k_2} \frac{\left[1-(k_1/k_2 \sin \Theta_{i1})^2\right]^{1/2}}{\cos \Theta_{i1}} \qquad \text{A (25)}$$

$$T^H_{21} = \frac{\overline{H}_{t1}}{\overline{H}_{i2}} = \frac{2}{1+Y^H_{21}} \qquad \text{A (26)}$$

$$R^H_{21} = \frac{\overline{H}_{r2}}{\overline{H}_{i2}} = \frac{1-Y^H_{21}}{1+Y^H_{21}} \qquad \text{A (27)}$$

$$R^H_{21} = T^H_{21} - 1 \qquad \text{A (28)}$$

$$Y^H_{21} = \frac{\mu_1 k_2}{\mu_2 k_1} \frac{\left[1-(k_2/k_1 \sin \Theta_{i2})^2\right]^{1/2}}{\cos \Theta_{i2}} \qquad \text{A (29)}$$

From these equations it is easily shown that

$$T^E_{12} = T^H_{21} \;;\; T^E_{21} = T^H_{12} \;;\; R^E_{12} = R^H_{21} \text{ and } R^E_{21} = R^H_{12} \qquad \text{A (30)}$$

If space 2 in figure 1 is a sheet between two semi-infinite media, it is necessary to determine the phase shift and attenuation of waves due to path length (l_2) in medium 2. A complex phase can be defined as

$$\phi = k_2 l_2 = \beta_2 l_2 - j\alpha_2 l_2 \qquad \text{A (31)}$$

When the boundaries are plane and parallel l_2 is related to the thickness (t_2) by

$$l_2 = \frac{t_2}{\sin \Theta_{t2}} \qquad \text{A (32)}$$

/3

where $\sin \Theta_{t2}$ can be complex by Snell's law (equation 2). Hence, 1_2 becomes a complex length. The complex impedance of the medium is

$$Z = \sqrt{\frac{\mu_2}{\epsilon_{T2}}} = 377\sqrt{\frac{|\mu_{r2}| e^{j\Theta\mu}}{\epsilon_{r2} - j\frac{\sigma_2}{\omega\epsilon_0}}} \qquad A\ (33)$$

the velocity of propagation is

$$v_{P2} = \frac{1}{\sqrt{\mu_2\epsilon_{T2}}} = \frac{1}{\sqrt{\left(|\mu_{r2}| e^{j\Theta\mu}\right)\left(\epsilon_{r2} - j\frac{\sigma_2}{\omega\epsilon_0}\right)}} \qquad A\ (34)$$

and the wavelength is

$$\lambda_2 = \frac{v_{P2}}{f} \qquad A\ (35)$$

We are now in such a position that we can calculate the transmission and reflection coefficients at each interface and the phase and attenuation in each space. Hence, the reflection and transmission coefficient of a multi-layered sandwich can be calculated by using successive (ABCD) matrix multiplication, as shown by Born and Wolfe[4] by assuming a positive going electromagnetic wave to be incident on the face of the sandwich and a negative going wave incident on the opposite face. Then, the TM mode (vertical polarization) wave matrix equation can be written for a multi-layer sandwich in matrix form as follows.

$$\begin{bmatrix} \overline{H}_{i1} \\ \overline{H}_{r1} \end{bmatrix} = \frac{1}{T_{12}^H} \begin{bmatrix} 1e^{jo} & -R_{21}^H \\ R_{12}^H & T_{12}^H T_{21}^H - R_{12}^H R_{21}^H \end{bmatrix} \begin{bmatrix} e^{\alpha_2 l_2} e^{j\beta_2 l_2} & o \\ o & e^{-\alpha_2 l_2} e^{-j\beta_2 l_2} \end{bmatrix} \times$$

$$\frac{1}{T_{23}^H} \begin{bmatrix} 1e^{jo} & -R_{32}^H \\ R_{23}^H & T_{23}^H T_{32}^H - R_{23}^H R_{32}^H \end{bmatrix} \begin{bmatrix} e^{\alpha_3 l_3} e^{j\beta_3 l_3} & o \\ o & e^{-\alpha_3 l_3} e^{-j\beta_3 l_3} \end{bmatrix} \times$$

$$\cdots \times \frac{1}{T_{nm}} \begin{bmatrix} 1e^{jo} & -R_{mn}^H \\ R_{nm}^H & T_{nm}^H T_{mn}^H - R_{nm}^H R_{mn}^H \end{bmatrix} \begin{bmatrix} \overline{H}_{tm} \\ \overline{H}_{im} \end{bmatrix} \qquad ; n = m-1$$

$$\qquad \qquad A\ (36)$$

14

where l_n is the complex path length in the n-th space. All other quantities have been defined above. For the TE mode (horizontal polarization), T_{nm}^E, T_{mn}^E, R_{nm}^E, R_{mn}^E, coefficients can be obtained from equation (30).

For a three-layer sandwich, this matrix in a short from is as follows

$$\begin{bmatrix} \overline{H}_{i1} \\ \overline{H}_{r1} \end{bmatrix} = \frac{1}{\Pi T_{nm}^H} \, [a]^H [b]^H [c]^H [d]^H [e]^H [f]^H [g]^H \begin{bmatrix} \overline{H}_{tm} \\ \overline{H}_{im} \end{bmatrix}$$

$$= \frac{1}{\Pi T_{nm}^H} \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix} \begin{bmatrix} \overline{H}_{tm} \\ \overline{H}_{im} \end{bmatrix}$$

A (37)

Setting $\overline{H}_{im} = 0$, then

$$R_{1m}^H = \frac{\overline{H}_{r1}}{\overline{H}_{i1}} = \frac{h_3^H}{h_1^H}$$

A (38)

and

$$T_{1m}^H = \frac{\overline{H}_{tm}}{\overline{H}_{i1}} = \frac{\Pi T_{nm}^H}{h_1^H}$$

A (39)

Additional layers can be added as desired with appropriate symbols for the overall (h) matrix. The E-mode (horizontal polarization) coefficients can be calculated likewise by substituting the T^E, R^E, quantities for the T^H, R^H, quantities in (36). Then,

$$\begin{bmatrix} \overline{E}_{i1} \\ \overline{E}_{r1} \end{bmatrix} = \frac{1}{\Pi T_{nm}^E} \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix}^E \begin{bmatrix} \overline{E}_{tm} \\ \overline{E}_{im} \end{bmatrix}$$

A (40)

where

$$R_{1m}^E = \frac{\overline{E}_{r1}}{\overline{E}_{i1}} = \frac{h_3^E}{h_1^E}$$

A (41)

and

$$T_{1m}^E = \frac{\overline{E}_{tm}}{\overline{E}_{i1}} = \frac{\pi T_{nm}^E}{h_1^E}$$

A (42)

VISUAL AID REQUIREMENTS

Name of Speaker_____Harris,_____Orville_____R._____
                       Last   First   Middle Initial

Title of Paper__The Strange Behavior of Electromagnetic Waves in Conducting__

Magneto-dielectric Media._____

_____

/X/   Vugraph Equipment

/ /   Slide Projector 35mm or 2" x 2"

/ /   Movie Projector 16mm

/ /   Opaque Projector

/ /   Blackboard

/ /   Electric Pointer

NOTE: If additional equipment is needed please state below.

_____

_____

_____

Incl 7

Title of Paper: The Strange Behavior of Electromagnetic Waves in

Conducting Magneto-dielectric Media

Dr. Orville R. Harris,
Name & Organization of Principal Author: U.S. Army Foreign Science and

Technology Center, 220 Seventh Street, N.E., Charlottesville, Virginia 22901

1. (X)   The abstract is UNCLASSIFIED. (Only unclassified abstracts will be accepted.)

2. (X)   The paper is UNCLASSIFIED.

3. ( )   The paper is CLASSIFIED. The classification is _____ .
This paper contains technological information or proprietary data which can be released to representatives of NATO and/or ABCA countries. (No paper classified above SECRET may be submitted.)

4.   The paper and abstract should be reviewed by ODDR&E ( ) Yes (X) No

N.A.
_____
Name, Title & Organization of Security Reviewer (Print or Type)

N.A.
_____
Signature of Security Reviewer - Date

5. (X)   This paper and abstract have been reviewed for content and editorial accuracy and are ready for presentation, publication and distribution. (No additional editing of this manuscript will be made prior to publication.)

Dr. Bertram B. Smith, Jr., Physical Scientist, U.S. Army Foreign Science
_____
Name, Title & Organization of Editorial Reviewer (Print or Type)
and Technology Center, 220 Seventh Street, N.E., Charlottesville, Virginia 22901

Bertram B. Smith 8 Mar 78
_____
Signature of Editorial Reviewer - Date

ARO Form 58
11 Aug 77

Inclosure 2

CULTIVATION OF HUMAN MALARIA PARASITES (U)

J. D. HAYNES, MAJ MC, *J. D. CHULAY, MAJ MC, C. L. DIGGS,
COL MC, R. E. DESJARDINS, MAJ MC, and CYNTHIA HALL, SP5
WALTER REED ARMY INSTITUTE OF RESEARCH
WASHINGTON, D. C. 20012

Malaria is a disease caused by protozoa of the genus Plasmo-
dium, the blood stage of which causes fever, anemia, and other
symptoms in infected humans. In every major armed conflict invol-
ving the United States, malaria has been a major cause of medical
disability. Of the four species responsible for disease in man,
Plasmodium falciparum is associated with the most illness. Because
malaria is still widely distributed throughout Africa, Asia, Oceania,
and Latin America, it remains the single most important potential
medical threat to military operations. Since the discovery nearly 100
years ago of the causitive agent of malaria, efforts at control of this
protozoan and its insect vectors have met with only limited success.
Recent experience with increasing resistance of mosquitos to insecti-
cides have made it clear that vector control is unlikely to eliminate
this disease. Similarly, chemotherapy has failed in large areas of
the world where P. falciparum has developed resistance to chloro-
quine and other drugs. Even in areas where drug resistance is not
yet a problem, the difficulties in maintaining strict compliance with
prophylactic regimens have resulted in failure to fully prevent dis-
ease.

Immunization would offer a logical means of protecting troops
from acquiring malaria when entering endemic areas, but progress
towards development of such immunization proceedures has been
hindered by a number of problems. Natural immunity to P. falci-
parum develops slowly and is incomplete. No completely satisfactory
animal model exists, and until recently malaria parasites could not
be cultivated in vitro. These limitations made it impossible to pre-
pare and test antigens which might potentially induce immunity. The
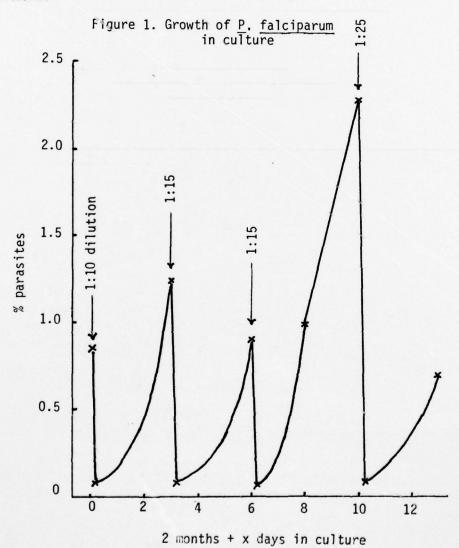
recent developments in our laboratory and at the Rockefeller University of methods for the continuous long-term culture of Plasmodium falciparum (1,2) make available a tool which can be used for studying pharmacologic, immunologic, and other biologic aspects of the host-parasite relationship. Antigens obtained from culture may eventually be used in preparing a vaccine against malaria.

The culture methods used involve mixing together infected and uninfected erythrocytes in buffered tissue culture medium, supplemented with human serum in a low oxygen, carbon dioxide-containing atmosphere, and incubating at 37 C. The infected erythrocytes can be obtained from humans, chimpanzees, or owl monkeys, and may be used immediately after collection, following cryopreservation, or from cultures. Uninfected erythrocytes can be from similar sources, although human erythrocytes stored using routine blood banking techniques are most convenient. Tissue culture medium can be either medium 199 (1) or RPMI 1640 (2), buffered with a bicarbonate-TES* or bicarbonate-HEPES* mixture. Human serum or plasma are equally effective at supplementing the medium, as long as it is obtained from a donor whose blood type is compatible with the erythrocytes used. The carbon dioxide content of the atmosphere can vary from 3 to 7%, adjusted to balance the bicarbonate in the medium to a pH of approximately 7.35 at 37 C. The oxygen content of the atmosphere can vary from 1 to 18%, but normal atmospheric oxygen (approximately 21%) is toxic for the parasite. The culture mixtures may be placed in tissue culture flasks which are individually flushed with the appropriate gas mixture, in microtiter plates which are placed in gas-tight boxes, or in petri dishes which are placed into a candle jar.

The current method which we use for maintenance of routine cultures consists of 5 ml samples with a 6% hematocrit in 25 sq cm tissue culture flasks containing M 199 buffered with 15 mM TES and 27 mM sodium bicarbonate, plus 10% heat inactivated human fresh frozen plasma. The flasks are flushed with a 5% carbon dioxide, 5% oxygen, 90% nitrogen mixture, sealed, and incubated at 37 C. Four ml of culture medium is changed daily, and samples for thin smears are made every 1 to 2 days. When parasitemia exceeds 1%, dilutions with fresh human erythrocytes are made to reduce the parasitemia to approximately 0.1%. Using this method, a 4 to 6-fold multiplication rate during each 48-hour cycle is obtained (figure 1). If the cultures are further diluted to 0.01%, fresh erythrocytes need not

*TES is N-tris-(hydroxymethyl)methyl-2-aminoethanesulfonic acid,
 HEPES is N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid.

be added for 7 days or more. By changing the medium more frequently, or by reducing the hematocrit and adding an extract of erythrocytes, higher parasitemias can be obtained. When these two methods were combined, parasitemias as high as 33% were achieved. In addition to morphological evaluation of growth, the incorporation of a radioisotope (tritiated hypoxanthine) can be used as a rapid and sensitive method of comparing parasite growth under different conditions.

Figure 1. Growth of P. falciparum in culture

The in vitro cultivation system has found a wide range of application. When erythrocytes from different species of animals are used as uninfected target cells, only those from species which are susceptible to infection with P. falciparum in vivo are susceptible to in vitro invasion (table 1). Differences in erythrocyte receptors for different species of plasmodia were found by evaluating in vitro reinvasion of test erythrocytes by Plasmodium knowlesi , a simian parasite capable of infecting rhesus monkey and most humans, or P. falciparum, which is unable to infect rhesus monkeys. Humans with blood type Fy(a-b-), also known as Duffy negative, cannot be infected by P. knowlesi (3). The results of in vitro reinvasion corresponded directly with the known in vivo susceptibility, suggesting that the two parasites recognize different erythrocyte receptors (table 2).

Table 1.   In vitro growth of P. falciparum in erythrocytes

test erythrocytes	% parasites day 1	% parasites day 3	multiplication rate/48 hr
human	.37	2.30	6.2
chimpanzee	.53	3.33	6.3
rhesus	.30	.13	.4
guinea pig	.55	.18	.3

(Adapted from reference 1)

Table 2.   Differential invasion by P. falciparum and P. knowlesi

test erythrocytes	Parasites invading test erythrocytes (% of control human erythrocytes)	
	P. falciparum	P. knowlesi
human A+	100	100
human Fy(a-b-)	109	1
rhesus	1	277

(Adapted from reference 3)

Further evidence supporting this concept was obtained when normal human erythrocytes were treated with different enzymes prior to attempts at in vitro infection (table 3). Chymotrypsin treatment markedly inhibited invasion by P. knowlesi, but had no effect on the growth of P. falciparum. Conversely either neuraminidase or

trypsin treatment inhibited reinvasion by P. falciparum without affect-
ing the growth of P. knowlesi. More knowledge about the parasite
surface molecules responsible for parasite-erythrocyte interactions
may aid in their purification. These molecules could then be used as
antigens in a vaccine which might elicit anti-parasite antibodies acting
to block the interaction of parasites with erythrocytes, thereby inter-
rupting the parasite life cycle and preventing disease.

Table 3. Invasion of enzyme treated human erythrocytes

enzyme treatment	Parasites invading treated erythrocytes (% of untreated control erythrocytes)	
	P. falciparum	P. knowlesi
chymotrypsin	102	10
neuraminidase	46	109
trypsin	22	102

(Adapted from reference 3)

P. falciparum grown in vitro can be used to evaluate antimala-
rial drugs. If serial 2-fold dilutions of drug are made in the wells of
a microtiter plate and aliquots of a culture are added to these, incu-
bated for 24 to 48 hours, and then pulsed with tritiated hypoxanthine
overnight, the effect of the drug on parasite growth can be assessed
by morphology or by measuring incorporation of tritium into polynuc-
leotides. One culture flask supplies enough parasites for 300 micro-
titer wells, which is sufficient to test in duplicate 7 different concen-
trations each of 18 different drugs. Semiautomated harvesting and
counting devices yield raw data within a few hours. Using regression
analysis, an estimate of the ED-50 (the dose at which there is 50%
inhibition of parasite growth) can be obtained. When a strain of
P. falciparum known to be sensitive to treatment with chloroquine
in humans was tested and compared with a known chloroquine-
resistant strain, a marked difference in the dose-response curves
was apparent (figure 2). When mefloquine, a drug known to be
effective for treating infections with either strain, was tested,
the ED-50 levels were similar (table 4)(4).

5

Figure 2.
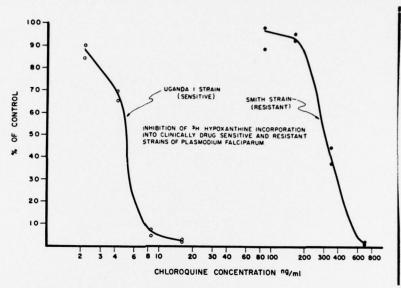
Table 4. Chloroquine sensitive and resistant P. falciparum strains evaluated in vitro

Drug	Concentration of drug (ng/ml) causing 50% reduction in uptake of tritiated-hypoxanthine.	
	Uganda strain	Vietnam strain
Chloroquine	8.8	184
Mefloquine	4.5	5.9

(Adapted from reference 4)

Current antimalaria drug screening tests generally are performed in vivo using animal models. These models are cumbersome because of the expense and manhours required, and in some cases yield data not applicable to human malaria. The method described here has the advantages of reducing screening time and expense, and uses only human malaria parasites. A spin-off of this research is that this method could be used to generate important information on the mechanisms of action of antimalaria drugs.

HAYNES, CHULAY, DIGGS, DESJARDINS and HALL

Cultured parasites may also be used to obtain antigen for a variety of immunological and serological tests. For example we have used these cultures as a source of antigen to detect antibodies using an indirect fluorescent technique. Thin smears made from culture were dried overnight, dilutions of test serum added, the slide thoroughly washed, flourescein-conjugated anti-human immunoglobulin reacted with the slide, and the slide washed and examined by fluorescence microscopy. Control sera from patients with no history of malaria gave titers no higher than 10, whereas the reference positive serum and sera from Africans living in an area endemic for malaria gave titers of 320 or higher (5). Similar titers were seen using either of two strains of P. falciparum taken from cultures 2 to 5 months old (table 5). This and other serologic techniques using antigen obtained from culture can be used in epidemiological studies of malaria transmission. Our results also demonstrate the ability to harvest antigens potentially important in the immune response to malaria.

Table 5. Indirect fluorescence antibody test using two strains of cultured P. falciparum as antigen

serum	IFA titer using Camp strain	Z strain
reference pos.	320	320
African 1	320	320
African 2	5,120	1,280
control 1	10	10
control 2	10	10

(Adapted from reference 5)

The culture system described here provides methodology for a variety of studies on the biology, immunology, and pharmacology of malaria. In addition to studies of parasite interactions with erythrocytes, evaluation of chemotherapeutic agents in vitro, and provision of antigen for serologic tests, once antigens are obtained in larger quantity from culture and purified, it should be possible to evaluate the feasibility of immunoprophylaxis against malaria.

REFERENCES

1.   Haynes, J. D., C. L. Diggs, F. A. Hines, and R. E. Desjardins. 1976. Culture of human malaria parasites *Plasmodium falciparum*. Nature. 263:767.

2.   Trager, W., and J. B. Jensen. 1976. Human malaria parasites in continuous culture. Science 193:673.

3.   Miller, L. H., J. D. Haynes, F. M. McAuliffe, T. Shiroishi, J. R. Durocher, and M. H. McGinniss. 1977. Evidence for differences in erythrocyte surface receptors for the malarial parasites, *Plasmodium falciparum* and *Plasmodium knowlesi*. J. Exp. Med. 146:277.

4.   Desjardins, R. E., J. D. Haynes, J. D. Chulay, and C. J. Canfield. 1978. Quantitative assessment of antimalarial activity by an automated *in vitro* system. Fed. Proc., Fed. Am Soc. Exp. Biol. in press (abstract).

5.   Hall, C. L., J. D. Haynes, J. D. Chulay, and C. L. Diggs. 1978. Cultured *Plasmodium falciparum* used as antigen in a malaria indirect fluourescence antibody test. Am. J. Trop. Med. Hyg. in press.

HARDENING OF COUNTERMINE STRUCTURES

DAVID C. HEBERLEIN, PhD
US ARMY MOBILITY EQUIPMENT RESEARCH AND DEVELOPMENT COMMAND
FORT BELVOIR, VIRGINIA

INTRODUCTION

Mines have long been recognized and used as effective barriers to retard or restrict the advance of enemy forces. Mines contain a sensor for target acquisition and an explosive kill mechanism that is directed towards a vulnerable area of an acceptable target. Mine explosive kill mechanisms include blast damage to armored vehicle tracks, wheels and suspension systems, shrapnel damage to personnel or non-armored vehicles, and shape charge damage to the "belly" of armored vehicles. Mine terminal effects can be defeated through the use of high strength, light weight composite materials. Vehicle components and countermine structures can be made from composite materials that will retain their functionality after being exposed to blast loading or high velocity fragments. Although these components are damaged by the mine, the retention of functionality permits the completion of mission. It is in this sense that vehicles and countermine structures are hardened against mine blast and shrapnel damage.

Computer codes are being developed to evaluate mine blast and shrapnel damage to countermine structures and vehicles. These codes are presently predicting damage to traditional armor steels for which a wealth of experimental data exists. The composite literature, while rapidly expanding, is not adequate to establish the shock compression equations-of-state for these materials. The objective of this investigation is to provide an experimental base for constructing the shock compression equations-of-state for selected composite materials.

SHOCK COMPRESSION EQUATION-OF-STATE

The utility of computer codes to meaningfully predict structural response to shock deposition is largely determined by the accuracy and completeness of the data used to determine the shock compression equations-of-state. Because structures are both heated and compressed during shock loading it is necessary to determine the temperature dependence of all independent variables in order to predict the time dependent response of the structure. The approach chosen to determine the material coefficients, A and B, in the shock compression equation-of-state

$$P = A \left(\rho/\rho_0 - 1 \right) + B \left(\rho/\rho_0 - 1 \right)^2, \tag{1}$$

where ρ is the material mass density, also determines whether $A=A(T)$ and $B=B(T)$ can be obtained from experimental data.

One direct method to obtain the temperature dependence of the material coefficients is through the Gruneisen formulation of the Debye theory of solids. An approach suggested by Harris and Avrami[1] expands Equation 1 above in terms of V for small volume changes as:

$$PV = A' V_0 + B' \frac{V_0}{V} + C'. \tag{2}$$

The Gruneisen parameter, Γ, can be expressed in terms of the speed of sound, c_o, as

$$\Gamma = - \frac{v}{c_o} \left(\frac{\partial c_o}{\partial v} \right)_T \tag{3}$$

where c_o is defined in terms of the longitudinal sound velocity, c_1, and the transverse sound velocity, c_t, as

$$\frac{1}{c_o^3} = \frac{1}{c_1^3} + \frac{2}{c_t^3} \tag{4}$$

Equations 2 and 3 can be combined to give the speed of sound in terms of the shock pressure, P, to give

$$c_o(\rho) = \left(\frac{\partial P}{\partial \rho} \right)^{1/2} \tag{5}$$

From equation 5, the speed of sound can be expressed in terms of the material coefficients by

$$C_o (\rho) = \left(\frac{A}{\rho_o}\right)^{\frac{1}{2}} \left(1 + \frac{2B}{A} \left(\frac{\rho}{\rho_o} - 1\right)\right)^{\frac{1}{2}} , \tag{6}$$

which can be approximated to give:

$$C_o (\rho) \cong C_o \left\{ 1 + \frac{B}{A} \left(\frac{\rho}{\rho_o} - 1\right)\right\} . \tag{7}$$

For $(B/A) ((\rho/\rho_o) - 1) \ll 1$, equation 7 reduces to

$$\frac{v}{c} \frac{\partial C}{\partial v}_o \cong \frac{v}{c} \frac{\partial}{\partial v} \left\{ C_o \frac{B}{A} \left(\frac{v_o}{v} - 1\right)\right\} \tag{8}$$

$$= - \frac{v}{C_o} \left\{ C_o \frac{B}{A} \frac{v_o}{v^2} \right\} , \tag{9}$$

giving the Gruneisen parameter at $V = V_o$ as

$$\Gamma = \frac{B}{A} . \tag{10}$$

The importance of this result is that the material coefficients for the shock compression equation-of-state can be established from fundamental thermodynamic properties of the materials themselves. The temperature dependence of the Gruneisen parameter is given by the thermal expansion coefficient, α, the specific heat, C_v, and the isothermal compressibility, K_T, as

$$\Gamma(T) = \frac{\alpha(T) \ V(T)}{C_V(T) \ K_T (T)} . \tag{11}$$

Thus, the temperature dependence of the ratio of the material coefficients of the shock compression equation-of-state may be determined directly through measurements of the thermal expansion coefficient, specific heat and isothermal compressibility. This paper reports experimental data of relative volume changes, thermal expansion coefficient, and specific heat as functions of temperature for selected composite materials over the temperature interval 240 to 370 K. These data when combined with sound velocity data are used to construct the time dependent shock compression equation-of-state for each material. The formulation of the computer code requires that three independent equations for shock pressure,

changes in internal energy, and the elastic pressure are expressed in terms of the material coefficients and the material mass density such that specification of initial conditions and the shock conditions defines the shock deposition process.

APPARATUS AND TECHNIQUES

The apparatus for measuring the length changes in the composite samples is shown in Figure 1. The changes in sample length were observed as changes in capacitance. The capacitor was of the three-terminal type. The area of the fixed, guarded plate was 6.41 cm^2. The capacitor gap was nominally 0.020 cm, resulting in a capacitance of approximately 20 pF. The capacitance was measured to approximately 1 part in 10^5, giving a length resolution of 10^{-7}cm. Sample lengths were typically 0.254 cm. A calibrated platinum resistance-thermometer was situated in a well in the sample platform. The stated calibration was checked both at the ice point and against the vapor pressure of liquid nitrogen between 64 and 77 K. The calibration points for the resistance R and temperature T were used in the relationship,

$$R = R_o \left(1 + \xi T + \eta T^2 \right), \tag{12}$$

to determine values for the constants ξ and η. Values for these constants were found that represented the stated calibration to an rms deviation of 0.03K over the temperature interval 230 to 370 K. Temperature changes were detected as resistance changes using a Hewlett Packard 3490A Multimeter. This instrument uses a four terminal measurement technique detecting changes as small as 0.001 ohm. The stated accuracy of the resistance measurement is 0.0001 percent. During the course of the experiments, the Hewlett Packard 3490A Multimeter and the Hewlett Packard 4270A Automatic Capacitance Bridge were used in the remote, addressed mode for acquiring data and transmitting it to a Hewlett Packard 9830A Calculator for data storage and display. The sample chamber was surrounded by a can into which helium exchange gas was introduced. A vacuum space thermally isolated this can from a surrounding liquid bath. A heater was wound on the exchange gas can to maintain the entire sample chamber assembly shown in Figure 1 at the desired temperature. The data acquisition system was programmed to maintain each temperature for a total of thirty measurements at each temperature. The capacitance-temperature data were then averaged before being stored in the memory of the HP 9830A Calculator.

TUNGSTEN
THERMOMETER

SAMPLE

PLATINUM
THERMOMETER

GLASS TO
METAL SEALS

Figure 2. Sample chamber for heat capacity measurements.

E

F G

H I

J

A

B

C

D
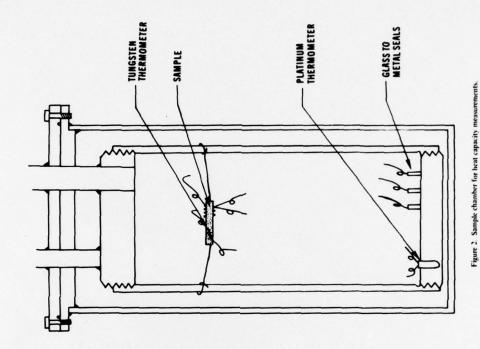
Figure 1. Sample chamber. A — lead to fixed capacitor plate, B — lead to active capacitor plate. C — platinum resistance-thermometer, D — bellows, E — adjustable bolts for fixed capacitor plate, F — concentric secondary capacitor plate, G — guard ring, H — sample, I — nylon posts, J — sample platform.

HEBERLEIN

The sample chamber used for the heat capacity measurements is shown in Figure 2. The sample chamber was surrounded by a vacuum can which was used as an adiabatic heat shield about the sample during the heat pulse measurements. A 100-ohm heater was wound on the outside of this inner vacuum can. The inner vacuum can was thermally isolated from the bath temperature by an outer vacuum can. During temperature calibration, the inner vacuum chamber was filled with helium exchange gas. This insured that the sample temperature was the same as that of the sample chamber assembly. The composite samples were suspended by two 0.002 cm diameter nylon wires from the walls of the sample chamber. The thermometer and heater leads to the sample were thermally anchored to glass-to-metal seals, which were soldered into the base of the copper sample chamber. Copper wires, 0.002 cm in diameter, were used as leads from the glass-to-metal seals to the sample to minimize the heat leak from the sample chamber to the sample. The sample heater was a 0.004 cm diameter constantin wire noninductively wound on the outside of the sample. A four terminal circuit was used to measure the current and voltage across the sample heater. The sample heater current was measured in terms of the voltage drop across a 1-kilohm resistor maintained in an external oil bath. The voltage across the 1-kilohm resistor and the voltage across the heater were measured with a Hewlett Packard 3490A Multimeter. The time duration of the heat pulse was detected with a Hewlett-Packard 5328A Frequency Counter. The stated acurracy of these measurements was better than 0.001 percent for the ranges which were used. The primary thermometer used in these measurements was a calibrated platinum resistance-thermometer as described for the relative length change measurements. The platinum resistance thermometer was situated in a well in the base of the sample platform shown in Figure 2.

Because the size of the samples was generally small (typically less than 3 grams), the mass of the sample thermometer was a major factor in maintaining the absolute accuracy of the heat capacity measurements. The use of standard resistance thermometry was precluded by the size and weight of commercially available platinum, germanium, and carbon resistors. The choice of the sample thermometer was also based on the nominal resistance of the temperature sensor, its temperature sensitivity, and its reproducibility on thermal cycling. The sample thermometer chosen consisted of a 5 cm piece of 0.0004 cm tungsten wire wound non-inductively about the sample. The wire was obtained in the form of a 100 ft (30.3 m) spool from Sigman Cohn Corporation. The tungsten wire was attached to the composite samples using a fine resin spray. This method of attachment proved to provide good thermal contact between the thermometer

and the sample. The room temperature resistance of the tungsten
resistance-thermometer was typically 100 ohms. The total resistance
of the tungsten resistance thermometer changed by a factor of 1.64
between 240 and 370 K. Changes in the resistance of the tungsten
resistance-thermometer were detected using a Hewlett Packard 3490A
Multimeter. The 3490A Multimeter uses a bridge principle for measure-
ment such that the nominal heat supplied to the temperature sensor
was less than 1 picowatt. In order to maintain a high precision of
temperature accuracy for the small temperature excursions of the sam-
ple during and after the application of the heat pulse, a function
was determined that gave the tungsten wire resistance in terms of the
temperature. The fit of the tungsten resistance over the temperature
interval 240 to 370 K had an rms deviation equal to or better than
that for the primary platinum thermometer. The function used to
determine the temperature from the resistance of the tungsten wire
was of the form:

$$ T = \sum_{i=0}^{i=4} A_i R^{(\text{slope})i} \tag{13} $$

Using this function, the 9830A Calculator could quickly and accurately
calculate sample temperatures.

The inner vacuum chamber was evacuated during the heat
pulse measurements. The inner vacuum can and the copper sample
chamber were maintained at a constant temperature during the heat
pulse measurements and, thereby, served as adiabatic heat shields.
The temperature of the adiabatic shield and the sample were maintained
at the same temperature for a period of 60 seconds before a heat
pulse sequence was initiated. After equilibrium between the sample
temperature and the sample chamber assembly had been maintained for
a period of 60 seconds, the sample temperature was allowed to drift
cold for a period of 30 seconds. This temperature drift before and
after the application of the heat pulse was used to correct for the
true temperature excursion during the application of a heat pulse.
The heat pulse duration was typically 15 seconds during which time
a Hewlett Packard 3490A Multimeter was used to make at least 15
separate readings of heater voltage, heater current and sample temper-
ature. After 15 seconds, the current to the heater was shut off
automatically and the temperature of the sample recorded as a function
of time for 200 seconds. The change in sample temperature was
determined by extrapolating the temperature drift both before and after
application of heat to the center of the heat pulse. The entire
process was controlled by a program entered into the memory of the
9830A Calculator. All data collection and analyses were performed

entirely within the Hewlett Packard 3050B Data Acquisition System. A permanent record of each heat pulse was recorded with a Hewlett Packard 9862A Thermal Printer.

DATA ANALYSIS

The raw data for the relative length change measurements consisted of a series of values of capacitance versus temperature at essentially zero pressure. The capacitance C for a three-terminal capacitor using a guard ring is related to the gap length l_g by

$$C = \frac{\pi \mathcal{E} r^2}{l_g} + \frac{\pi \mathcal{E} rw}{l_g + 0.22w} \left(1 + \frac{w}{2r} \right) \tag{14}$$

where r is the radius of the plate with the guard ring, $\mathcal{E}$ is the permittivity, and w the half-width of the spacing between the inside radius of the guard ring and the radius of the plate. The experimental geometry $r \gg w$ made the effect of the last term in Equation 14 on dC/C less than 0.3 percent and was neglected. Changes in sample length l_s are found directly in terms of C from Equation 14 because $l_s + l_g$ = constant. The linear expansion coefficient was obtained by differentiating Equation 14 with respect to temperature at constant pressure to give:

$$\left(\frac{dl_g}{dT} \right)_P = \frac{\pi \mathcal{E} r^2}{C^2} \left(\frac{dC}{dT} \right)_P + \frac{2\pi \mathcal{E} r}{C} \left(\frac{dr}{dT} \right)_P . \tag{15}$$

The last term in Equation 15 was also neglected because the radial expansion of the copper plates was small; i.e.

$$\frac{1}{C} \left(\frac{dC}{dT} \right)_P \gg 10^4 \left(\frac{1}{r} \right) \left(\frac{dr}{dT} \right)_P \tag{16}$$

Since the gap length change is equal and opposite to the sample length change, $dl_s = -dl_g$, the linear expansion coefficient is given by

$$\alpha_\ell = \frac{\pi \mathcal{E} r^2}{l_s C^2} \left(\frac{dC}{dT} \right)_P \tag{17}$$

Because measurements were taken along three perpendicular axes for each sample, the volume expansion coefficient was obtained by adding the three linear expansion coefficients for each sample.

The specific heat of a material is defined as the heat capacity normalized per mass, volume, or mole. For a solid maintained

at essentially zero pressure, the specific heat at constant pressure, C_p is given by:

$$C_p = \frac{1}{m} \frac{(\Delta Q)}{(\Delta T)} = \frac{1}{m} \frac{V_h I_h \delta t}{T} \tag{18}$$

where m is the mass of the sample, ΔQ the total amount of heat added to the sample, V_h, the voltage across the heater, I_h, the current in the heater, and δt the total time the heater was on. The heat leak from the heater leads to the sample was corrected through the use of a three-terminal measurement technique of heat supplied to the heater. In this technique two matched leads are connected from the glass-to-metal seals in the base of the chamber assembly to one side of the sample heater. A third matched lead extends from the other side of the sample heater to a third glass-to-metal seal thermally anchored to the sample platform assembly. In this manner, exactly one-half of the ohmic heating occurring in these lead wires is auto-matically counted in the measurement of power supplied to the sample heater. The other half of the ohmic heating is not counted because it is absorbed by the sample platform assembly. The additional heat capacity of the sample heater and thermometer was much smaller than that of the composite samples, C sample $\geq 10^2$ ($C_{htr} + C_{therm}$), and corrections to the calculated specific heat values for the masses of the heater and sample thermometer were also neglected.

COMPOSITE SAMPLES

The composite samples were obtained from Dr. I. Daniel, IIT Research Institute, Chicago, Illinois. The samples selected were representative of different types of composites as well as similar materials with different plies. The composite selection included cross-ply E-Glass, cross-ply S-Glass, unidirectional S-Glass, boron/aluminum 1928-9, boron/aluminum 1928-11 and KEVLAR. The c-axis was taken as that axis perpendicular to the plies. The a and b axes were identified as spatially orthogonal to the c axis.

The resin composites absorb moisture from the air over prolonged periods of explosure. Before the samples were mounted in the sample chamber, each sample was held in an oven at 373 K for a period of approximately 1 week. After the samples were mounted in the sample chamber, the samples were heated to 380 K, and a vacuum established about the sample. The heated samples were maintained in vacuo for a period of day before experiments were initiated.

RESULTS

A typical plot of relative length change versus temperature
is shown in Figure 3 for the c-axis of a KEVLAR/ERLA 4617 sample.
The linear expansion coefficient can be obtained through a point-by-
point differentiation of the data shown in Figure 3. The scatter
inherent in this technique is approximately $\pm$ 5%. In order to handle
the numerous calculations for which these data will be used, a
different technique was used to calculate the linear expansion coeffi-
cient. An analytic function was calculated to fit the relative
changes in length. Typically, 8 to 10 points from the relative length
change versus temperature curves were used to find an analytic function
of the form:

$$\frac{\Delta \ell_s(T)}{\ell_s(T)} = \sum_{i=0}^{i=N} A_i \, T^i \tag{19}$$

The linear expansion coefficient is then found by differentiating the
analytic function found from Equation 19 with respect to temperature.
The linear expansion coefficient corresponding to the data shown in
Figure 3 for relative length changes along the c-axis in KEVLAR is
shown in Figure 4. The volumetric expansion coefficient is then ob-
tained by adding the linear expansion coefficients for the three
independent spatial axes. The volumetric expansion found for samples
of boron/aluminum 1928-9 is shown in Figure 5.

The specific heat data were taken at approximately 0.5 K
temperature increments from 240 to 370 K. The raw specific heat data
for cross-ply S-Glass epoxy composite is shown as a function of
temperature in Figure 6. To facilitate handling of the data, an
analytic function of the form

$$C_P(T) = \sum_{i=0}^{i=9} C_i \, T^i \tag{20}$$

was fit to ten experimental specific heat versus temperature data
points. From the smoothed curves obtained for the volumetric expans-
ion coefficient and those obtained for the specific heat from Equation
20, it is possible to show the temperature dependence of Γ/C_o^2, where
C_o is the speed of sound in the solid material. The Gruneisen para-
meter is given by

$$\Gamma(T) = \frac{\alpha V}{K_T C_v} = \frac{\alpha C_o^2}{C_v} \, . \tag{21}$$
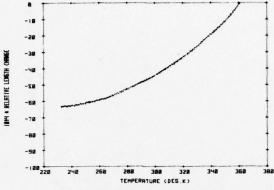
Figure 3. Relative Length Changes in KEVLAR/ERLA 4617 (c-axis)
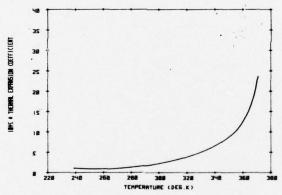
Figure 4. Linear Expansion Coefficient for KEVLAR/ERLA 4617 (c-axis)
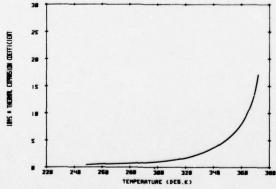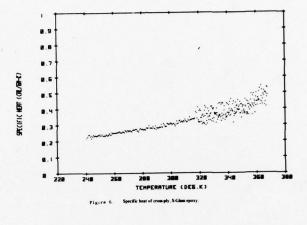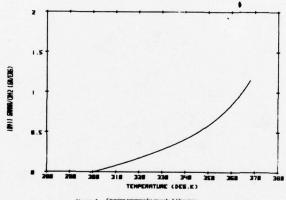
Figure 5. Volumetric Expansion of Boron/Aluminum 1928-11
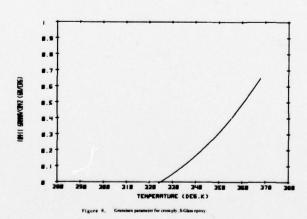
The temperature dependence of the quantity $10^{11}\ \Gamma/C_o{}^2$ (g/erg) is shown
for samples of cross-ply E-Glass and cross-ply S-Glass epoxy com-
posites in Figures 7 and 8 respectively. Because the experimental
data that leads to the formation of Figures 7 and 8 is too voluminous
to present herein, the differences in their temperature dependence is
summarized. The heat capacities of the S-Glass composites were found
to be higher than those measured for the E-Glass composites. The
thermal expansion coefficients, while the same order of magnitude
for both type of Glass epoxy composites, exhibited a much stronger
temperature dependence for the S-Glass composites. In fact, both
S-Glass composites exhibited a negative volumetric expansion coeffi-
cient for temperatures below approximately 325 K. Because the iso-
thermal compressibility is generally slowly varying with temperature
over the temperature interval of interest, it can be seen from Equa-
tion 21 that the quantity α/C_v shown in Figures 7 and 8 is a good
approximation to the temperature dependence of the Gruneisen parameter.
Although neither isothermal compressibility nor sound velocity
measurements are presently available over this temperature interval,
it is reasonable to assume that the strong temperature dependence
of α/C_v shown in Figures 7 and 8 and similiarly found for the other
composites measured in this investigation indicates that the use of
only the room temperature value of the Gruneisen parameter in shock
compression equations-of-state will lead to erroneous or uninterpret-
able predictions of the time dependent response of a composite
material to shock loading.

APPLICATIONS TO DEVELOPMENTAL WORK

Composite materials present certain physical properties
that make them promising candidates for use in hardening vehicles and
countermine equipment to mine blast and shrapnel damage. The heat
capacities measured for the composite materials are of the same order
of magnitude as that found for strong metals. The composite materials
are capable of absorbing large quantities of energy from shock
loading. The thermal expansion of composite materials is generally
slightly higher than that used in armor materials. The composite
materials are considerably lighter than armor materials while retain-
ing strength properties equal to that found in armor steels. The
composites can be formed with the reenforcement plies perpendicular
to the direction of anticipated loading. Armor steels are essentially
homogeneous solids and critical loading results in catastrophic fail-
ure to an entire structure. Composite materials, while exhibiting
local failure to blast loading, retain their structural form and thus
their functionality. In addition, the composite material absorbs

HEBERLEIN

Figure 6. Specific heat of cross-ply, S-Glass epoxy.

Figure 7. Gruneisen parameter for cross-ply, E-Glass epoxy

Figure 8. Gruneisen parameter for cross-ply, S-Glass epoxy.

large amounts of the shock energy and does not pass this energy through it to some other weaker link in the structure.

For development work, it is necessary to consider not only the physical properties of the composite materials but also the cost of acquisition and tooling workable structures. Shown in Figure 9 is a cylindrical form built principally of low cost KEVLAR. This particular KEVLAR structure has been subjected to the blast and shrapnel from ten M-21 mines (M-21 mines contain 10 lb of H-6 explosive). The structure survived these repeated shots at distances ranging from 2 to 6 feet at the solid angle found to have the highest density of metal mine fragments. This particular structure was developed to protect a countermine coil, used to clear magnetic influence mines, from blast and shrapnel damage. Shown in Figure 10 are two fully hardened countermine coils to be used in field testing against live mines. The composite material used to harden the coil to blast and fragment damage amounts to less than 40% of the weight of the total coil structure.

CONCLUSIONS

Future measurements of the sound velocity of these composites will define the temperature dependence of the Gruneisen parameter and the temperature dependence of the ratio of the material coefficients in the shock compression equations-of-state. These data will permit the development of predictive codes that can trace the time evolution of shock deposition in various structures. The high specific heat values found for composites indicates that these materials can be used to absorb large quantities of energy and thereby preclude the passage of this energy to weaker links in the structure. The temperature dependence of the quantity α/C_v indicates a strong dependence of the Gruneisen parameter, Γ, on temperature and consequently the importance of defining $B(T)/A(T)$ for composites before development of predictive codes.

REFERENCES:

P. Harris and L. Avrami, "Some Physics of the Gruneisen Parameter", Picatinny Arsenal Technical Report No. 4423, September 1972.
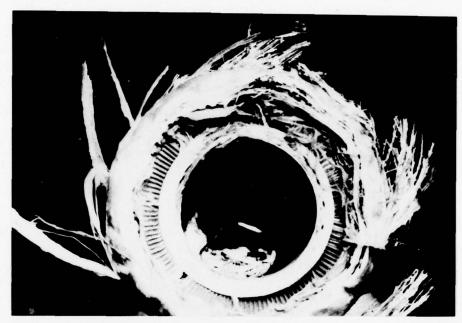
Fig. 9. Countermine Coil Form Subjected to Repeated Blast and
Shrapnel of M-21 Mines
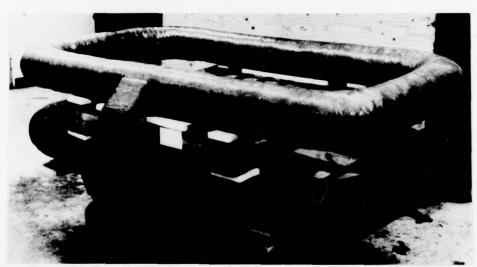
Fig. 10. Two Completely Hardened Countermine Coil Forms

HIGH-SPEED ELECTRON-BEAM LITHOGRAPHIC RESISTS FOR
MICRON AND SUBMICRON INTEGRATED CIRCUITS (U)

JOHN N. HELBERT, PhD, CHARLES C. COOK, JR., MR.,
AND EDWARD H. POINDEXTER, PhD
USA ELECTRONICS TECHNOLOGY AND DEVICES LABORATORY (ERADCOM)
FORT MONMOUTH, NEW JERSEY 07703

INTRODUCTION

Next-generation RADAR and ELINT systems are planned to provide
the field commander with comprehensive intelligence on the disposition
of enemy weapons and electronics equipment. Ultra-compact signal proc-
essors with unprecedented capabilities are the heart of these systems.
Fabrication of the required high-density integrated circuits (IC's),
with elements in the micron to submicron range, is beyond the resolu-
tion limit of state-of-the-art optical photolithography. Electrons
with 10-20 keV energies can be accurately focused to beam diameters
much less than a micron. Electron-beam lithography (EBL) meets the
projected resolution requirements, and is expected to be a key tech-
nology for the production of sophisticated new digital communications
systems for the Army.

Like the photolithographic IC production technique, EBL is also
totally dependent upon organic polymer resists for pattern delineation.
In EBL processing, a 5,000-10,000 Å resist film is spun onto the mono-
lithic IC substrate. IC patterns are then written into the resist with
the finely focused electron beam. E-beam irradiation decreases or in-
creases the resist molecular weight, which, in turn, increases or de-
creases the dissolution rate relative to the unirradiated portion of
the resist layer. When the irradiated resist pattern dissolves at a
faster rate, an indented image is formed, and the resist is termed a
positive acting resist (see Figure 1). If the opposite occurs, the
resist is termed to be negative in function (see Figure 2).

Economical EBL production hinges upon the resist sensitivity,
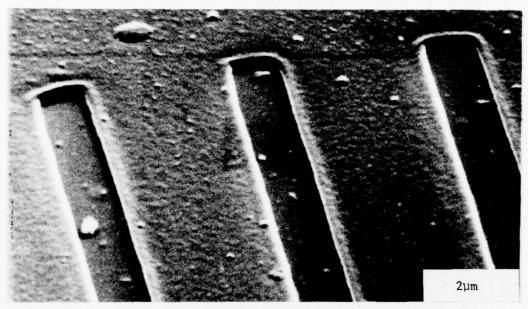
Figure 1. Scanning electron micrograph of developed electron-beam etched lines in poly(methyl methacrylate).
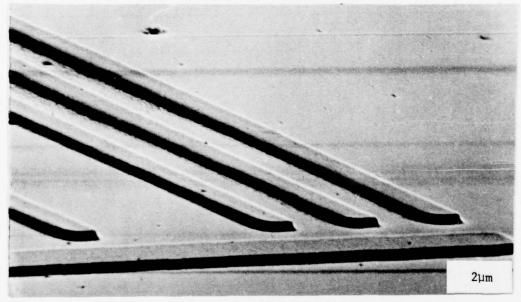
Figure 2. Scanning electron micrograph of developed electron-beam etched lines in poly(alpha-chloroacrylonitrile).

which governs the electron writing time per wafer. The objective of our work is a set of sensitive resists with varying properties for the production of next-generation high-performance military IC devices. Early efforts in this technological area have been largely based upon the availability of commercial polymers. In this work, we apply basic radiation-chemistry principles, which were more or less disregarded previously, to devise and develop high-merit polymer resists.

RADIATION EFFECTS AND MEASUREMENT TECHNIQUES

When a polymer resist is irradiated in-vacuo with 10-20 keV electrons (or equivalent gamma-irradiation), both chemical and physical changes occur. First, the polymer molecular weight changes due to bond-scission or crosslinking.

Changes in number-averaged molecular weight, $\bar{M}_n$, obey the equation: (1)

$$\bar{M}_n^{-1} = \bar{M}_{n,o}^{-1} + \left[G_s - G_x\right] D/100 \; N_A \tag{1}$$

where G_s and G_x are the radiation scission and crosslinking yields per 100 eV absorbed dose, N_A Avogodro's number, and D the radiation dose. Similarly, the weight-averaged molecular weight, $\bar{M}_w$, follows the equation: (2)

$$\bar{M}_w^{-1} = \bar{M}_{w,o}^{-1} + \left[G_s - 4G_x\right] D/200 \; N_A \tag{2}$$

Thus if $(\bar{M}_n^{-1})$ and $(\bar{M}_w^{-1})$ are followed as a function of dose, G_s and G_x may be obtained from the slopes of the two independent determinations. Values of $\bar{M}_n$ are usually obtained by membrane osmometry, while $\bar{M}_w$ is obtained by either gel permeation chromatography (3) or light scattering techniques (3). For G_s predominant the resist is positive; for G_x, negative.

When polymers undergo radiation-induced main-chain scissions, the formation of free radicals is a consequence. Under certain conditions some of these radicals are stable enough to be observed by electron paramagnetic resonance techniques (4). As in scission and crosslinking radiation processes, radiation-scission radical values, G(rads), can be measured. Polymers with high G(rads) are also found to have high $G_s - G_x$ or G_s (5).

The last technique, but most important to electron resist researchers, is the direct e-beam sensitivity (Q) method. Q is the total absorbed e-beam dose in coulombs/cm^2 and is measured by scanning

electron microscopy techniques. For a positive resist, Q is defined as the dose at which 20% of the unexposed resist is removed during development for complete dissolution of the irradiated beam pattern area. This is also equivalent to a dissolution rate ratio, $d_{irradiated}/d_{unirradiated}$ (6), of 5. For a negative resist, Q is defined as the dose at which a 50% thickness of irradiated resist remains following total development dissolution of the unexposed resist portion.

All of the techniques discussed above are capable of measuring the radiation degradation or crosslinking susceptibility of a polymer resist system. The independent G values have been found to be well correlated (7).

POLYMER RESIST STRATEGY

Before directly testing a polymer as a predetermined type of a resist, an arduous and time-consuming task, the radiation behavior of that polymer should be reviewed if it has been previously investigated. If it hasn't been studied at the basic level, it is probably time-efficient to do one or more of the measurements described in Radiation Effects and Measurement Techniques, than to proceed to direct testing. This is, in fact, how resist research at ET&DL is carried out. The basis for several of the e-beam resists utilized today can be found in the basic polymer radiation chemistry papers published between 1958 and 1973.

Wall (8) and Miller et al (9) observed very early that vinyl polymers of the form $-(CH_2-C(X)Y)-$, where X nor Y is hydrogen, degrade preferentially when subjected to ionizing radiation in vacuo. Polymethyl methacrylate (PMMA) $-(CH_2-C(CH_3)CO_2CH_3)-$, a high-resolution positive e-beam resist, is an example; PMMA has a G_s value of 1.3 (10), a G_x value of zero (11), a $G(CO_2CH_3)$ of 1.0 (12), and Q is 5×10^{-5} C/cm^2 (13). Although PMMA is capable of high resolution, its sensitivity Q is too low for commerical utilization in EBL.

Our approach to the development of a more sensitive positive resist is to retain the basic known degrading structural unit $-(CH_2-C(X)Y)-$, but to replace the X or Y group (e.g., CH_3 for PMMA). By substituting electron-withdrawing or electronegative groups at X or Y, we hoped to weaken the energy required for degradation and to increase resist sensitivity. This effect is found experimentally for molecular organics. The C-C bond energy in $CH_3-CO-CO-CH_3$, for example, is 60 kcal/mole, compared to 83 kcal/mole for H_3C-CH_3; the electronegative oxygen substituents on the carbons reduce the C-C bond energy by 28%.

Following this polymer selection criterion, four basic vinyl polymers, poly(methacrylonitrile)(PMCN) $\{CH_2-C(CH_3) \ CN\}$, poly(methyl alpha-chloroacrylate)(PMCA) $\{CH_2-C(Cl)CO_2CH_3\}$, copolymer P(MCA-co-MMA), and copolymer P(MCN-co-MCA), have been studied at the basic level. PMCN represents a substitution of the Y substituent with an electron-withdrawing cyano (CN) group and PMCA represents a substitution of the methyl (CH$_3$) group at X with electronegative chlorine (Cl). In addition, poly(isobutyl methacrylate $\{CH_2-C(CH_3) \ CO_2CH(CH_3)_2\}$ was studied to determine the effect of the bulky ester alkyl group; poly(alpha-hydroxy isobutyric acid) $\{C(CH_3)_2-COO\}$ was selected and studied to determine if the polyester unit $\{COO\}$ combined with the well-known degrading quaternary carbon unit $\{C(CH_3)_2\}$, could enhance G_s or Q over those values of PMMA reference. Results of these basic radiation studies are found in Table I.

PMCA and P(MCA-co-MMA) are seen to be 2-4X more susceptible to radiation degradation than PMMA (14). Higher radiation degradation susceptibility is also observed for PMCN (15) and P(MCN-co-MCA). Direct e-beam testing is warranted for all these systems. In contrast, poly(alpha-hydroxy isobutyric acid) (16) and poly(isobutyl methacrylate (17) were found to be less susceptible to radiation degradation than PMMA, and e-beam testing was not warranted.

Similar basic radiation studies of positive resists poly(butene-1-sulfone) (18), poly(isobutylene) (10), and poly(alpha-methyl styrene) (20) were the foundations for several of the results listed in Table I.

E-BEAM RESIST TESTING RESULTS

Positive Resists

Negative e-beam resists possess lower intrinsic resolution capability than positive resists; therefore, the ETDL resist program has dealt mainly with positive resist development -- except for one system (see Negative Resist Section). In addition, negative resists exhibit stronger tendencies to swell during development processing, which tends to lead to bridging of written features and resist distortions (21). Utilization of negative resists for direct-write e-beam lithography in the submicron domain will probably be excluded.

Following our basic radiation studies where higher G values were observed for PMCA versus PMMA reference (14), we determined that PMCA functioned as a more sensitive positive resist (22) than PMMA. Informal collaboration with Texas Instruments (TI) group working under ECOM contract #DAAB07-75-C-1297 led to further testing of PMCA; TI

TABLE I. G values and e-beam sensitivities for known positive resist polymers.

Positive Resist System	G_s-G_x	G_s	G_x	G(rads)	Q, C/cm²	Lab
Poly(alpha-methyl styrene)	--	0.3	--	0.05	1×10^{-4}	BTL[a]
poly(alpha-hydroxy isobutyric acid)	0.30	0.37	0.07	--	not studied low G_s value	ETDL[b]
poly(isobutyl methacrylate)	1.1	1.1	0	--	not studied low G_s value	ETDL;HRC[c]
PMMA	1.3	1.3	0	1.9	5×10^{-5}	IBM[d]
poly(isobutylene)	--	4.0	--	2.5	2×10^{-5}	BTL;ETDL
poly(methyl alpha-chloroacrylate) (MCA)	5.4[e](2.3)[f]	6.0[e](3.2)[f]	0.7[e](0.5)[f]	5.7	1×10^{-5} $3-6\times10^{6}$[g]	ETDL;HRC TI[h]
poly(MCA-co-MMA)	2-3	3.1	0.06	3.1	7×10^{-6}	ETDL;HRC
poly(methacrylontrile) (PMCN)	3.3	--	--	--	$4-7\times10^{-6}$	ETDL
poly(MCN-co-MCA)	4-5	--	--	--	1.2×10^{-5}	ETDL
poly(MCN-co-MMA)	--	2-3	--	--	$3-5\times10^{-6}$	BTL
poly(butene-1-sulfone)	--	12	--	3	$1-4\times10^{-6}$	BTL

a BTL - Bell Labs

b ETDL - Electronics Technology & Devices Laboratory (ERADCOM)

c HRC - Honeywell Research Center

d IBM - International Business Machines

e G value observed at low doses

f G value observed at high doses

g TI data from ETDL report #DAAB07-75-C-1297

h TI - Texas Instruments

i $\bar{M}_n$ was only 1.2×10^5 for this resist; an increase of three or four in $\bar{M}_n$, however, will increase Q to $3-4\times10^{-6}$ C/cm².

6
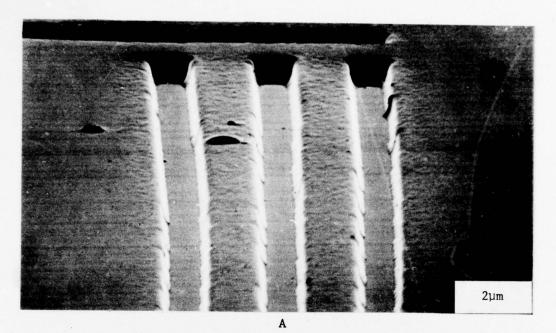
found the resist to be even more sensitive than reported by ETDL (see Table I). This resist is currently undergoing advanced device production testing under ERADCOM contract #MMT-76-C-8105. Although this polymer is less sensitive than the Bell Labs licensed poly(butene-1-sulfone) resist (see Table I), it possesses superior shelf-life stability, substrate adhesion, and ion-etch "dry-processing" compatibility. PMCA does, however, crosslink to some extent as determined by the G_x=0.5-0.7 value from the basic studies, but this does not seriously impair its positive resist behavior. This deficiency can be overcome, at the expense of G_s, by utilization of the P(MCA-co-MMA) copolymer system. The basic study results of Table I show that G_s is decreased about twofold while G_x is decreased about tenfold. This explains why Lai et al (23) of Honeywell were able to observe better Q values for their P(MCA-co-MMA) copolymer systems than for their MCA homopolymer resist sample. Our earlier basic studies with P(MCA-co-MMA) helped lay the groundwork for their results and they referenced our study (24).

The G values of Table I for PMCN and P(MCN-co-MCA) dictated direct e-beam testing for these systems. They are found to function as sensitive positive resists. See Figure 3 for sample e-beam etched patterns for these systems. Preliminary measurements indicate, that Q is $4-8\times10^{-6}$ C/cm^2 for PMCN and $1-2\times10^{-5}$ C/cm^2 for the copolymer (70% MCA - 30% MCN). It should be emphasized here that Q is strongly a function of development processing, and that these processes require a considerable amount of research to produce optimization. Following our lead into the MCN polymer resist formulation (25), Bell Labs independently reported (26) that P(MCN-co-MMA) functions as a very sensitive positive e-beam resist (see Table I).

Negative Resists

Positive resists act as high-resolution negative resists when overexposed. The doses required, however, are usually 1.5-2 orders of magnitude higher than usual. Poly(alpha-chloroacrylonitrile) (PACAN) $-(CH_2-C(Cl)CN)-$, an early ETDL test positive system, changes resist behavior at significantly lower electron charge densities. The Q value at 10 keV is found in Table II. Although PACAN is not as sensitive as the Bell Labs P(GMA-co-EA) system, it is more sensitive than the TI polystyrene resist (27) and is capable of high resolution (see Figure 2).

PACAN must crosslink at the alpha-chlorine site (i.e., the X site). Chlorine removal (i.e., C-Cl bond cleavage) evidence for irradiated poly(vinyl chloride) $-(CH_2-C(Cl)H)-$, and Penton $-(CH_2-C(CH_2Cl)_2-CH_2-O)-$ is well established (28,29). Although the

7
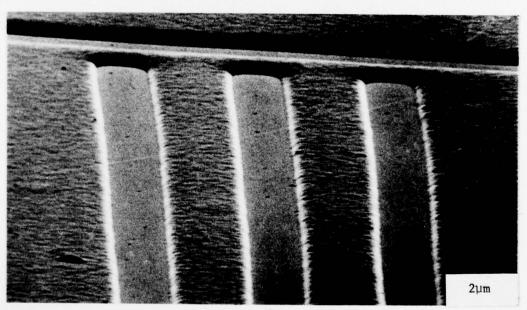
Figure 3. Scanning electron micrographs of developed electron-beam etched lines in (A) poly(methacrylonitrile) and (B) the copolymer poly(methacrylonitrile-co-methyl alpha-chloroacrylate).

C-Cl bond is weaker than the C-C bond in these polymers, there is evidence that the chlorine in Penton is cleaved off by a dissociative electron capture reaction (29) which is thermodynamically favorable (30). Negative resist behavior for PACAN may also be the result of a low glass transition temperature, t_g. It is notable that structurally similar PMMA, PMCA and PMCN all have t_g values above 100 C and the crosslinking reactions for these polymers are strongly suppressed.

TABLE 11. Negative e-beam resist sensitivities for four representative systems.

Resist System	Q at 10 kV, C/Cm^2	Lab
Polystyrene	2×10^{-4}	TI[a]
PACAN	5×10^{-5}	ETDL[b]
P(GMA-co-S)	5×10^{-6}	HRC[c]; BTL[d]
P(GMA-co-EA)(COP)	4×10^{-7}	BTL[d]

[a] Texas Instruments

[b] Electronics Technology and Devices Lab (ERADCOM)

[c] Honeywell Corporate Research Center

[d] Bell Labs

RESIST SUMMARY

Only two e-beam resists, one negative (COP) and one positive (poly(butene-1-sulfone)), are commercially available. These two resists offer high sensitivity, but are beset with undesirable properties as well. The best positive resist, excluding sensitivity, is still standard PMMA. The ETDL positive test systems, PMCA and PMCN, are not optimized to the extent of the Bell Labs commercial resists, but preliminary results are very encouraging and advanced testing warranted. PMCA and P(MCA-co-MMA) resists have higher ion-etch compatibilities and possess potential as dry-process submicron production resists.

HELBERT, COOK, POINDEXTER

Much in-depth research is still needed, both at the basic and production levels, to insure attainment of a suite of resists with desirable and diversified properties suitable for specific large scale integration devices.

REFERENCES

1. M. Dole, in "The Radiation Chemistry of Macromolecules," Vol. II, M. Dole, Ed., Academic Press, New York, 1973.

2. R.W. Kilb, J. Phys. Chem. 63, 1838 (1959).

3. P.E. Slade, "Polymer Molecular Weights," Marcel Dekkar, New York, 1975.

4. D. Campbell, J. Polym. Sci.: D Series, 4, 91 (1970).

5. A.R. Shultz, in "Chemical Reactions of Polymers," Vol. 19, E.M. Fettes, Ed., Interscience, New York, 1964, p. 657.

6. J.N. Helbert, C.F. Cook, Jr., E.H. Poindexter, and B.E. Wagner, Polym. Eng. and Sci. 17, 414 (1977).

7. L.F. Thompson and M.J. Bowden, J. Electrochem. Soc. 120, 1722 (1973).

8. L.A. Wall, J. Polym. Sci. 17, 141 (1955).

9. A.A. Miller, E.J. Lawton, J.S. Balwit, J. Polym. Sci. 14, 503 (1954).

10. E. Gipstein, A.C. Ouano, D.E. Johnson, and O.U. Need, III, IBM J. Res. and Dev., 21, 143 (1977).

11. A.R. Shultz, P.I. Roth, and G.B. Rathman, J. Polym. Sci. 22, 495 (1956).

12. A. Todd, J. Polym. Sci. 42, 223 (1960).

13. M. Hatzakis, J. Electrochem. Soc. 116, 1033 (1969).

14. J.N. Helbert, P.J. Caplan, and E.H. Poindexter, J. Appl. Polym. Sci. 21, 797 (1977).

15. J.N. Helbert, E.H. Poindexter, G.A. Stahl, R.C. Chen, and C.U. Pittman, Jr., J. Polym. Sci.: Chem. Ed., 0000 (1978).

16. C.U. Pittman, M. Iqbal, R.C. Chen, and J.N. Helbert, Polymer Letters, (to be published).

17. J.H. Lai and J.N. Helbert, Macromolecules, 0000 (1978).

18. J.R. Brown and J.H. O'Donnell, Macromolecules 5, 109 (1972).

19. J. Danon, J. Chim. Phys. 62, 281 (1965).

20. A.M. Kotliar, J. Appl. Polym. Sci. 2, 134 (1959).

21. R.D. Heidenreich and G.W. Kammlott, Fourth Photopolymer Conference, Society of Plastics Engineers, Ellenville, N.Y., p. 95 (1976).

22. J.N. Helbert, C.F. Cook, Jr., and E.H. Poindexter, J. Electrochem. Soc. 124, 158 (1977).

23. J.H. Lai, L.T. Shepard, R.P. Ulmer, and C.H. Griep, Forth Photopolymer Conference, Society of Plastics Engieers, Ellenville, N.Y., p. 259 (1976).

24. J.N. Helbert, B.E. Wagner, P.J. Caplan, and E.H. Poindexter, J. Appl. Polym. Sci. 19, 1201 (1975).

25. J.N. Helbert, E.H. Poindexter, G.A. Stahl, R.C. Chen, and C.U. Pittman, Jr., Polymer Preprints 18, No. 1, 772 (1977).

26. L.E. Stillwagon, E.M. Doerries, L.F. Thompson, and M.J. Bowden, Coatings and Plastics Preprints 37, No. 2, 38 (1977).

27. T.L. Brewer, 6th Int. Conf. Electron Ion Beam Sci. Technol, R. Bakish, Ed., 71 (1974).

28. A.A. Miller, J. Phys. Chem. 63, 1755 (1959).

29. Y.J. Chung, D.R. Squire, and V. Stannett, J. Macromol. Sci.-Chem., A8, 1403 (1974).

VELOCITY MEASUREMENTS ABOUT A NACA 0012 AIRFOIL WITH
A LASER VELOCIMETER

*DANNY R. HOAD, WARREN H. YOUNG, JR., PhD,
STRUCTURES LAB, USA RESEARCH & TECHNOLOGY LABS (AVRADCOM)
JAMES F. MEYERS
NASA LANGLEY RESEARCH CENTER
HAMPTON, VA 23665

SUMMARY

A laser velocimeter was installed in the Langley 4- by 7-meter
low-speed (V/STOL) tunnel to measure the velocity field about a wing
with a NACA 0012 airfoil section. These measurements were compared at
two low angles of attack (0°, 4.15°) with a two-dimensional viscous-flow
prediction program. At 0°, the comparison provided confidence in the
effectiveness and accuracy of the laser velocimeter. At 4.15°, the data
indicated that a small laminar separation bubble with oscillating shear
layer probably existed. The unique capability of the laser velocimeter
in measuring absolute flow magnitude and direction without prior know-
ledge of general flow direction was demonstrated in the complex separated
reverse flows over the wing at an angle of attack of 19.4°.

INTRODUCTION

The laser velocimeter (LV) is a non-intrusive velocity measure-
ment instrument. Its application has been demonstrated in difficult
situations where conventional devices either cannot survive or their
presence would severely influence the measurement desired. Reference 1
describes the application of the LV in determining the velocity charac-
teristics within turbine stages of an engine assembly where pressure
probe blockage was found to affect the velocity measurement significantly.
In very low velocity situations with large turbulent structure, the LV

has been found to be very effective (see refs. 2 and 3). In this
situation, measurement by conventional devices would have been very
cumbersome with large rapid traversing mechanisms which would have
inherently influenced the velocity field. In many situations, the
direction of the velocity at the desired measurement location is not
known. In this case, conventional measurement techniques require
careful measurements of flow direction before precise velocity measure-
ments could be obtained. Often the researcher is faced with the problem
of obtaining velocity measurements where flow direction is unknown and
insertion of conventional devices will cause damage to the instrument
or the model. The determination of the in-plane velocity characteris-
tics of a helicopter rotor system provides the most comprehensive use
of the unique capabilities of the LV (ref. 4).

An LV system installed in the 4- by 7-meter low-speed (V/STOL)
tunnel at Langley was designed to accomplish two objectives: (1) to
demonstrate and verify the use of the LV in this facility by comparison
with theoretical techniques, and (2) to determine the flow characteris-
tics over a stalled three-dimensional wing.

SYMBOLS

The axes used for this investigation are presented in figure 1.
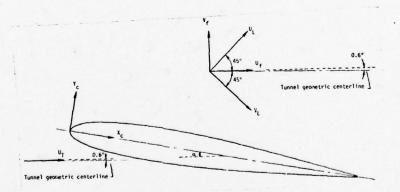
Figure 1. - Sketch of axis system used including
directions of velocity components computed.

The velocity measurement position was referenced to the airfoil chord-
line, and the velocity measurement magnitude was referenced to the free-

stream direction. The units for the physical quantities defined in this paper are the International System of Units. Most quantities were measured in this system; however, some were measured in the U.S. Customary Units and converted by using factors given in reference 5.

c   wing chord, 0.3048 m

N   number of velocity measurements in one ensemble

N_i number of velocity measurements in i^{th} histogram interval as percent of N

U_f local velocity component, parallel to direction of free-stream velocity, m/sec

U_L local velocity component in LV optics coordinate system, 45° above tunnel centerline (see fig. 1)

U_R local total velocity, $\sqrt{U_f^2 + V_f^2}$, m/sec

U_T free-stream velocity determined from pitot-static probe, m/sec

V_f local velocity component, perpendicular to direction of free-stream velocity, m/sec

V_L local velocity component in LV optics coordinate system, 45° below tunnel centerline (see fig. 1)

X_c, Y_c coordinate axis relative to wing chord

x_c distance downstream from airfoil leading edge along chord, m

y_c distance above wing chord, perpendicular to it, m

α wing angle of attack, deg

APPARATUS

A fringe-type LV optics system operating in the backscatter mode was used for these tests. A sketch of the optics system is presented in figure 2, and a photograph is presented in figure 3. A high-speed burst counter was used to measure the period of the high-frequency signal

contained in the burst from a particle traversing the sample volume. LV system control, data acquisition, and data reduction were handled by
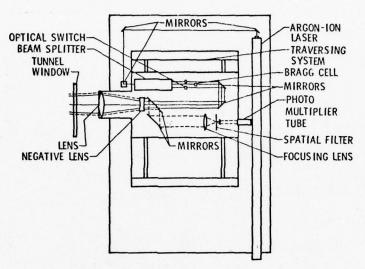
Figure 2. - Schematic of the LV optics.
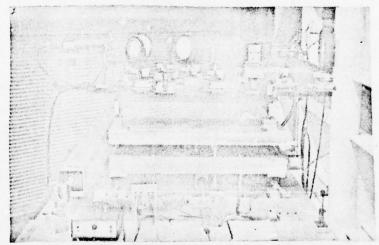
Figure 3. - Photograph of the LV optics.

a minicomputer. A complete description of the LV optical system, electronics system, and data acquisition and reduction is available in reference 6.

The model used in this investigation was a simple straight wing. It had a span of 2.438 m, a chord of 0.3048 m, and a NACA 0012 airfoil section. Velocity measurements were made at center span to obtain two-dimensional characteristics. The wing was supported by struts from the floor near the tunnel centerline with no balance measurements taken. The location of the strut mount to the wing was chosen as far outboard as structurally feasible to minimize flow disturbance at the wing centerline. A photograph of the model with crossing laser beams is presented in figure 4.
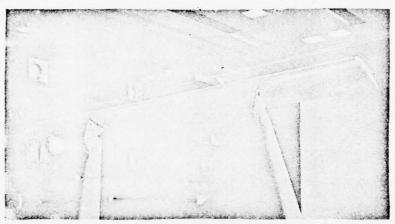
Figure 4. - Straight wing with NACA 0012 airfoil
section installed in V/STOL tunnel with
crossing laser beams.

This ivestigation was conducted in the Langley 4- by 7-meter low-speed (V/STOL) tunnel at a nominal free-stream Mach number of 0.15. The Reynolds number based on the wing chord was approximately 1×10^6. Local flow velocities were measured about the wing centerspan at three angles of attack: 0°, 4.15°, and 19.4°. The low angle-of-attack measurements were obtained to compare with a two-dimensional theoretical prediction technique (ref. 7), and the high angle-of-attack measurements were obtained to define the flow-field characteristics over a separated airfoil.

A pitot-static probe was mounted 2.5 m below and 1 m ahead of the the wing centerline to provide accurate reference for the free-stream tunnel dynamic pressure. A hygrometer was used to obtain wet bulb temperatures and, total temperature was measured in the settling chamber. Thus, the tunnel air density could be calculated, and with dynamic pressure measurements, the tunnel velocity could be accurately determined.

DISCUSSION

Velocity measurements at each measurement location were first reduced to histogram form. These data for the wing at $\alpha = 0°$ and $4.15°$ (relative to tunnel geometric centerline) along with a description of data reduction technique, histogram interpretation, and complete error analysis can be found in reference 6.

Free-stream velocity measurements were obtained with the LV in the clear tunnel at the location of the wing centerline. These data indicated an average upwash angle of $0.6°$ (relative to tunnel geometric centerline). The wing was installed with the chordline parallel to test-section centerline; therefore, the effective angle of attack was assumed to be $0.6°$. It is known, however, that the presence of the wing and struts change this angle, but the magnitude of this change is not known.

Prediction Technique

The external forces generated on a body in a fluid are manifested in the velocity distribution of the fluid about the body. In developing a prediction technique, the calculations at the surface of the body are verified with conventional pressure and force measurements. Reference 7 presents an excellent comparison with measured surface pressures for this viscous-flow prediction. Since the local surface pressures are computed from predicted local surface velocities, it is justifiable to question the validity of the predicted velocities away from the surface. The need to verify these prediction velocities is evident. The use of conventional probes near the surface raises questions about the accuracy of the measurement because of the interferences caused by the presence of the probe. It was determined that the LV was a device capable of measuring this flow field without inducing any interference since nothing was present in the field but the wing and light beams.

The theory for this prediction technique (ref. 7) involves an iterative procedure which first obtains an inviscid-flow solution for the basic airfoil. It computes a boundary-layer solution based on the inviscid-flow solution and constructs a modified airfoil by adding the boundary-layer displacement thickness to the original airfoil. It obtains the inviscid solution for the modified airfoil and repeats these steps until appropriate convergence criteria are satisfied. The field point velocities are then computed from the vorticity distribution along the modified airfoil.

Experiment-Theory Comparison

At $\alpha = 0°$. - Velocity vectors as measured by the LV for the
wing at $\alpha = 0°$ are presented in figure 5. Each velocity vector (arrow)

SCANS
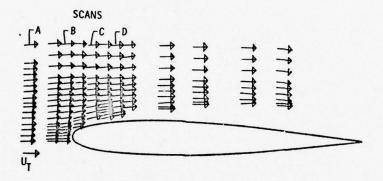
Figure 5. - Velocity vectors computed from
measurements over the wing at a geometric
angle of attack of 0°.

is an average of an ensemble of measurements taken over a short period
of time at the desired location. This arrow plot indicates the relative
location of the velocity measurements, magnitude and angle of the velo-
city vector. The velocity magnitude and direction is indicated by the
length and orientation of the arrow. The position of velocity measure-
ment is marked by the tail end of the arrow. The tunnel free-stream
magnitude and direction reference is provided in the lower left corner
of the figure.

The velocity measurements were obtained by positioning the sam-
ple volume at a desired chordwise station (x_c/c), and incrementing the
entire optics package downward along this chordwise station. This was
accomplished remotely and was completely controlled by the minicomputer.
Four of these series of measurements (scans A, B, C, and D from fig. 5)
are presented in more detail in figure 6. This figure presents a
comparison between LV measured velocities and the two-dimensional
viscous-flow prediction. The comparison is presented with the resultant
velocity nondimensionalized by tunnel free-stream velocity as a function
of the vertical position of the measurement nondimensionalized by wing
chord.
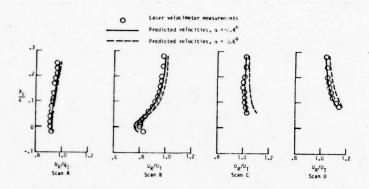
Figure 6. – Comparison of LV flow-field velocity
measurements with a two-dimensional viscous-flow
prediction. Wing geometric angle of attack = 0°.

The free-stream upwash angle without wing or supports was measured at
0.6° at this Mach number. Typically, flow angularity is affected by a
model's presence. It is normally determined by model upright and inver-
ted angle-of-attack ranges. Comparisons of balance data from these two
conditions provides the total flow angularity. It is very difficult to
obtain this type of measurement with discrete velocity measurements in
the presence of the model. Since no balance measurements were obtained
in this investigation, there is some uncertainty in the effective angle
of attack of the wing. Predicted velocities were calculated first using
the measured 0.6° tunnel flow angle without the wing. These are presen-
ted as dashed lines with maximum discrepancies on the order of 6 percent.
Calculations were repeated with a one degree shift in angle of attack to
provide an assessment of the effect of uncertainty in this measurement.
These calculations are presented as solid lines with $\alpha = -0.4°$ and
indicated better agreement with theory. It is obvious in these compari-
son that the precise measurement of these velocities depended on the
precise determination of the effective angle of attack of the wing. It
is justified to say, however, that these data provide a quantitative
and qualitative measure of the accuracy and acceptability of LV measure-
ments about a lifting surface.

At $\alpha = 4.15°$. – The velocity vectors as measured by the LV for
the wing at a geometric angle of attack of 4.15° are presented in
figure 7. Four of these series of measurements (scans A, B, C, and D
from fig. 7) are presented in more detail in figure 8. As before, this
presents a comparison between LV measured velocities and the two-
dimensional viscous-flow prediction for the four scans with large
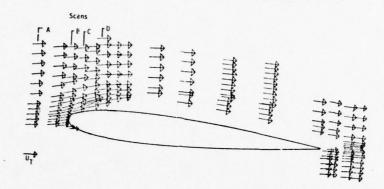
Figure 7. – Velocity vectors computed from measurements over the wing. Geometric angle of attack = 4.15°.
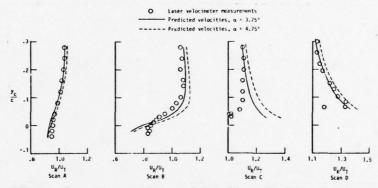
Figure 8. – Comparison of LV flow-field velocity measurements with a two-dimensional viscous-flow prediction. Wing geometric angle of attack = 4.15°.

velocity gradients. Predicted velocities using the measured 0.6° tunnel flow angle (without the wing) are presented as dashed lines. The agreement is not good particularly very near the leading edge (scans B and C). Based on the data presented for the wing at 0°, an adjusted flow angularity (-1.0°) was chosen to obtain an approximation of the effective angle of attack (3.75°). These data are presented as a solid line and show good agreement for scans A and D, but still poor agreement for scans B and C. The LV measured velocities for scan C particularly indicate a different shape in the velocity profile. This suggests that more than flow angularity uncertainty existed. The basic histogram data
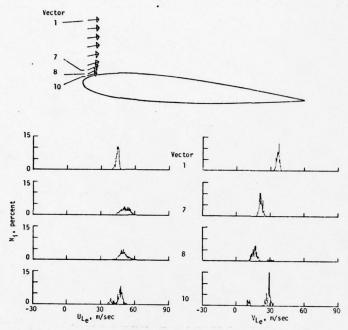
Figure 9. - Histogram data for scan C.
Wing geometric angle of attack = 4.15°.

for four of the velocity vectors within this scan are presented in figure 9. The local flow characteristics for these vectors are presented in histogram form for vectors 1, 7, 8 and 10.

A sketch of the wing cross-section with arrows indicating the position, direction, and relative magnitude of the mean velocity vectors is also provided in figure 9. A vector is an average of an ensemble of data acquired at the position desired. The histogram is a graphical representation of the variation of velocity measured over a time period. They are presented with N_i, percentage of that number of measurements within incremental velocity band, as a function of velocity. The basic LV coordinate system U_L component is presented on the left and the V_L component on the right. Interpretation of histogram information is provided in reference 6.

The histograms above the surface (vectors 1-7) are well defined Gaussian-type distributed velocity measurements. Very near the surface the histogram (vectors 8-10) has two peaks, which indicate two predominant velocity values. At this position and subsequent positions, the flow experiences an oscillation between the two values, sometimes with greater

tendency to be at or near one value than the other, but spends little
time between the two general values. This type of histogram has been
shown to be generated by the passage of a transverse vortex through the
measurement volume (ref. 8). Information presented in reference 9
indicated that this airfoil section should have a laminar separation
bubble near the leading edge at this angle of attack operating at this
Reynolds number. The most likely explanation of these double-peaked
histograms and the very poor agreement with theory was that there
existed a laminar separation bubble on the upper surface with a thin
oscillating shear layer. If the shear layer were steady, the double
peaks would probably not exist. As described in reference 6, the mea-
surements leading up to this point indicated variation in velocity and
angle. The position of the separation point, as measured in reference
9, is highly sensitive to slight wing angle-of-attack changes; there-
fore, it is possible that the separation point was moving with the
tunnel flow angle oscillation. This unsteadiness in the separation
point would trigger an oscillating shear layer.

The Airfoil at 19.4°

An arrow plot of the mean velocity field about the airfoil at
$\alpha = 19.4°$ is presented in figure 10. Each velocity vector is presented
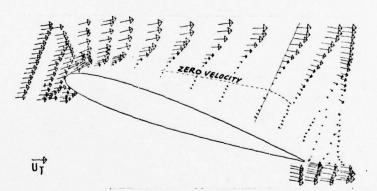
Figure 10. - Velocity measurements over the stalled
wing ($\alpha = 19.4°$).

with the length of the arrow indicating magnitude and direction of the
vector relative to the airfoil. It is obvious from this figure that the
airfoil is in a fully separated condition. The shear layer region
between the free-stream and the separated turbulent area over the wing
is broad, but easily discernable. The velocity field in the separated
region indicates the existence of a large recirculating eddy with

reverse flow near the airfoil surface. The velocity fluctuations within
the shear layer were large; however, it the reverse flow region, the
velocity fluctuations were smaller. A dashed line is provided indi-
cating the approximate location of zero velocity in the separated
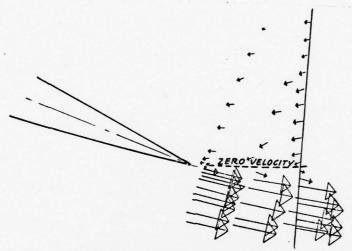region. At the trailing edge (see fig. 11), a very sharp shear layer

Figure 11. – Velocity measurements at the trailing
edge of the wing at   = 19.4°.

is evident with low, reversed flow velocities generated near the airfoil
upper surface and with nearly free-stream velocity from the lower surface.
The spatial distance across this shear layer is on the order of 0.005
x/c. The reverse flow in the wake region above the airfoil is also
evident.

These velocity measurements were obtained without prior know-
ledge of the direction of the flow at each measurement point. The LV
is unique in this capability unlike conventional probes which require
this information to reduce ambiguity primarily caused by alinement
requirements and support structure interference (ref. 10).

CONCLUSIONS

A laser velocimeter was installed in the Langley 4- by 7-meter
low-speed (V/STOL) tunnel to measure the velocity field about a straight
wing with a NACA 0012 airfoil section. The wing was installed at three
geometric angles of attack: 0°, 4.15°, 19.4°. This was done to provide

data at low angles of attack to compare with a well-accepted two-dimensional viscous-flow prediction program and at a high angle of attack to characterize the flow field in the separated region over a fully stalled wing.

The results of this investigation indicated that:

1. The laser velocimeter is an effective and accurate instrument for measuring the velocity field over a surface.

2. The precision of the laser velocimeter measurements for the wing at a geometric angle of attack of 0° depended on the precise determination of the effective angle of attack of the wing in the tunnel.

3. The data for the wing at a geometric angle of attack of 4.15° indicated that a laminar separation bubble probably existed with a thin oscillating shear layer.

4. The separated region over the wing in the fully stalled condition was well defined. Reverse flow measurements in this region demonstrated the unique capability of the laser velocimeter for measuring velocity magnitude and direction without prior knowledge of the flow direction.

5. The trailing-edge measurements at the highest angle of attack demonstrated the capability of the laser velocimeter in measuring the velocity characteristics across a very sharp velocity gradient.

REFERENCES

1. Goldman, Louis J.; Seasholtz, Richard G.; and McLallin, Kerry L.: Velocity Surveys in a Turbine Stator Annular-Cascade Facility Using Laser Doppler Techniques. NASA TN D-8269, 1976

2. Gartrell, Luther R.; and Jordon, Frank L., Jr.: Demonstration of Rapid-Scan Two-Dimensional Laser Velocimetry in the Langley Vortex Research Facility for Research in Aerial Applications. NASA TM 74081, August 1977

3. McLaughlin, Dennis K.: Laser Doppler Velocimeter Measurements in a Turbulent Jet Exiting Into a Cross Flow. Arnold Engineering & Development Center Rep. TR-71-262, January 1972

4. Biggers, James C.; and Orloff, Kenneth L.: Laser Velocimeter Measurements of the Helicopter Rotor-Induced Flow Field. J. Am. Hel. Soc., vol. 20, no. 1, Jan. 1975, pp. 2-10

5. Mechtly, E. A.: The International System of Units - Physical Constants and Conversion Factors (Second Revision). NASA SP-7012, 1973

6. Hoad, Danny R.; Meyers, James F.; Young, Warren H., Jr.; and Hepner, Timothy E.: Laser Velocimeter Survey About a NACA 0012 Wing at Low Angles of Attack. NASA TM 74040, 1978

7. Smetana, F. O.; Summey, D. C.; Smith, N. S.; and Carden, R. K.: Light Aircraft Lift, Drag, and Moment Prediction - A Review and Analysis. NASA CR 2523, 1975

8. Young, Warren H., Jr.; Meyers, James F.; and Hepner, Timothy E.: Laser Velocimeter Systems Analysis Applied to a Flow Survey Above a Stalled Wing. NASA TN D-8408, 1977

9. Gault, Donald E.: An Experimental Investigation of Regions of Separated Laminar Flow. NACA TN 3505, 1955.

10. Wentz, W. H., Jr.; and Seetharam, H. C.: Split Film Anemometer Measurements on an Airfoil with Turbulent Separated Flow. Fifth Biennial Symposium on Turbulence, Univ. of Missouri, Rolla, Missouri, Oct. 3-5, 1977, pp. I.3-1-I.3-8

CATASTROPHIC REACTION OF COMPARTMENTALIZED
AMMUNITION - CAUSES AND PREVENTIVE MEASURES

*PHILIP M. HOWE, Ph.D
ROBERT B. FREY, Ph.D
BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MD 21005

INTRODUCTION

While it is feasible to design a tank ammunition compartment
which will survive the detonation of a single warhead, the design of
a compartment which will survive the detonation of most or all of
the warheads and which falls within the weight and space constraints
imposed by the vehicle design is not currently possible; the detonation
of 40 warheads (the planned complement) will destroy compartment and
fighting vehicle. In previous research (1) we have shown that
catastrophic reaction of munitions can occur under conditions much less
strenuous than those required for classical shock initiation, which
has been studied extensively for bare charges. These catastrophic
reactions play an extremely important role in determining munition
vulnerability, and in the rapid propagation of explosion through
stacks of munitions. Typically, these catastrophic reactions take
place in the 100-700μsec time frame, consume essentially all of the
explosive, and may appear to be detonations to the observor interested
in assessing damage potential. To understand the mechanisms of
initiation of these reactions, and to devise preventive techniques
suitable for safe transportation and storage, and vulnerability reduc-
tion of armored fighting vehicles, we have undertaken analyses and a
variety of experiments. Pertinent results are summarized in this
paper.

Interround Communication and the Role of the Casing

When a munition such as a 155mm artillery shell is detonated, nearby munitions are subjected to multiple fragment impacts, airblast, and severe loading from the explosive products. Initiation of the target munitions can occur as a result of a single, massive, high velocity fragment, as a result of multiple fragment impacts occurring nearly simultaneously, or as a result of the severe blast loading delivered by fragments and explosive products. One might conjecture that some measure of protection would be provided the target munitions by heavy walled casings, but it must be remembered that a heavy walled protective casing of a target round implies massive lethal fragments when such a round serves as the donor.

Because there is a variety of parameters which may affect interround communication, a series of experiments was performed to establish a data base and to provide insight with respect to the pertinent mechanisms. Munitions from the inventory were used rather than specially designed test fixtures, because the former would provide a much needed practical data base and because analysis had shown that the variations in geometry from munition to munition could be accounted for and would not weaken the validity of the results. Each experiment involved three munitions placed collinearly upon a 2.5 cm steel witness plate (see Figure 1). The two outer munitions served as targets for the center warhead which was deliberately detonated via primacord embedded within some C-4 plastic explosive which filled the fuze cavity. (In some cases, target warheads contained fuzes. However, fuze presence did not change threshold response, and there was no evidence in any experiment to indicate that the fuze contributed to reaction of the target warheads). The donor wall thickness, diameter, and explosive content, and acceptor wall thickness, diameter, and explosive content were varied. Data were obtained for munitions containing composition B (60%$_w$ RDX, 40%$_w$ TNT) and composition A-3 (91% RDX 9% wax).

Figure 1. Schematic of typical test configuration for interround communication

The condition of the acceptor warheads and the witness plate was examined after each experiment to determine acceptor response. When a donor was detonated in the design mode, it always perforated the witness plate and this was taken as a crude but effective indicator of acceptor detonation. For each type of warhead pair, the separation distance between rounds was varied in accordance with standard quantal response techniques (2) to determine the propagation threshold. Separate tests were performed on inert loaded projectiles placed at separation distances at which the acceptor detonated. This permitted determination of the level of damage which would lead to violent reaction or detonation of the acceptor warheads. At the violent reaction threshold, each inert-loaded acceptor was severely deformed and failure of the warhead casing occurred. This provided an important clue with respect to the mechanism by which violent reaction occurs within the acceptor. All the data are consistent with a mechanism involving

. casing deformation
. compression of the explosive, generation of cracks within the explosive
. failure of the casing
. rapid extrusion of explosive through cracks in casing, causing ignition and rapid spread of reaction through the cracked explosive, with resultant catastrophic explosion.
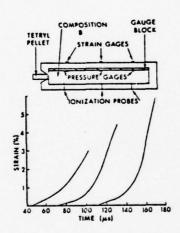
Figure 2. Schematic of apparatus used for mechanistic studies. Typical strain/time records at various gauge locations.

Additional experiments were conducted to explore some of the details of the initiation process. In one set of experiments, heavily confined composition B charges were fabricated, with internal manganin pressure gauges and externally mounted constantan strain gauges (see Figure 2). These charges were deliberately ignited, in order to permit observation of the development of violent reaction.

The detailed behavior of the charges was variable and strongly a function of geometry. Thus, for some experiments, a compression

wave propagated through the charge at a velocity of 2.0 - 2.5 mm/μsec. This wave was clearly not a shock, as the pressure gradient typically extended over a period of 10 - 30 μsec behind the wave front, to a peak pressure of 0.2 - 0.8 GPa. Generally speaking, strain records and stress records were similar. A plot of strain versus time at various gauge stations for such an experiment is shown in Figure 2. Ionization probes indicated reaction begins within a few microseconds of wave passage. Nonetheless, the pressures involved are too low to cause ignition by rapid compression of the explosive (3,4). (Note that, even with ignition, the reaction would not necessarily build up to violent reaction or detonation. In some instances, localized reaction occurs, and disrupts part of the charge without propagating to the rest.)

In other experiments, using larger diameter charges the pressure rose slowly but uniformly throughout the charge. After 200 - 600 μsec, a threshold was reached, at which point the pressure rose very rapidly and catastrophic reactions occurred. As in the previous experiments, the thresholds at various stations occurred sequentially and are associated with the arrival of a compression wave. In these latter experiments, the compression waves decayed as they propagated, and ionization probes responded erratically, indicating that low-level secondary ignition sources were developing at various locations. In both sets of experiments, large strains and strain rates were recorded coincident with the point at which the pressure transition points occurred.

At first we thought that charge deformation might be creating adiabatic shear bands which caused secondary ignition of the charge at points remote from the initial reaction. We analyzed the rate of temperature rise when two layers of explosive slip with respect to one another under pressure and with melting. The analysis indicates that sliding velocities of the order of $3X10^3$ cm/sec would be necessary for initiation at 0.1 GPa pressure. At lower sliding rates melting would suppress the temperature rise. In our experiments where the rate of shear deformation has been measured, such sliding velocities at shear bands could occur only if the shear bands were separated by distances of the order of a centimeter, which is unlikely. Therefore, it appears that the formation of shear bands does not explain the propagation of reaction in these experiments, although they could be responsible for the initiation of reaction in other circumstances.

An alternative explanation, consistent with the interround communication data and fragment impact data, is that homogeneous (on a macroscopic scale) deformation of the explosive does not cause secondary ignition, but that ignition results from casing failure and extrusion of the explosive into the cracks formed as the case opens.

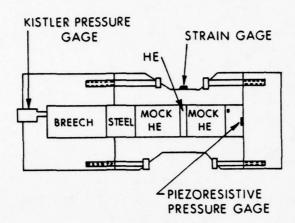An experiment was designed to test this hypothesis. It is shown schematically in Figure 3. Propellant was burned in the breech, thereby driving a plastic piston into the explosive, which was held in place by a deformable cylindrical container. The explosive was subjected to pressures and deformations similar to those of earlier experiments, but the deliberate ignition source was eliminated. In this experiment, ignition and violent reaction always occurred, but only after the metal case ruptured. We conclude that ignition and the development of violent reaction in confined charges is intimately connected with casing failure.

Figure 3. Apparatus used to show that ignition results from casing failure.

Huffington, in a parametric study of the response of thin shells to external blast loading, considered effects of geometry, loading, and material properties for fixed end cylinders subjected to a "frontal cosine" distribution of impulsive loading (3). The geometry is shown in Figure 4.

The shells were considered to be thin ($D/h < 1$) and Kirchhoff's hypothesis was applied (5). A mathematical formulation nonlinear in the equations of motion, the elasto-plastic stress strain relations, and the strain displacement relations was developed. The behavior of the solution was explored by varying non-dimensional ratios one

Figure 4. Frontal cosine impulsive loading for a fixed-ended cylinder (3).

at a time, holding others constant. For complete details, the reader is referred to the original paper. Of special interest, however, is the fact that both maximum and residual deformation are strong functions of a scaled impulse density, $\frac{i\,C_o}{Eh}$, (where i is the impulse/ area C_o the speed of sound in the casing, h the casing wall thickness, and E Young's modulus), and the fact that these functions depend only weakly on dimensionless ratios such as length/diameter, (L/D) and case thickness/diameter (h/D); (see Figure 5).

Figure 5. Casing response vs impulse intensity.

This is particularly important; it permits the identification of a critical deformation for casing failure with a unique value of the scaled impulse density delivered to the target, and the threshold inter-round communication distance can be obtained by equating the scaled impulse density to some critical value

$$\pi_3 = \frac{i\,C_o}{Eh} = \pi_{crit}.$$ This

critical value must be obtained from experiment. For fragmenting munitions contained within typical arrays such as tank ammunition compartments or pallets, the fireball of the donor munition envelopes the vicinal munitions.

Thus, both fragments and explosive products contribute to the impulse density delivered to the acceptors. We calculated values of the average areal fragment momentum according to the relation, applicable for cylindrical changes:

$$\frac{mV}{A} = \frac{\rho\,(r_o^2 - r_i^2)}{2R}\,V$$

where ρ is the donor casing material density

r_i is the donor casing inside radius

r_o is the donor casing external radius

v is the average fragment velocity, calculated using Gurney Formulae.

R   is defined by Figure 1.

The total specific kinetic energy, E_{HE}, is given by

$$E_{HE} = \frac{\overline{V}_{HE}^2}{2}$$

where $\overline{V}$ is the average product speed. E_{HE} is proportional to the Gurney energy, E^*, so

$$\overline{V} \propto \sqrt{2E^*}$$

and the scaled areal impulse ratio delivered by the explosive varies according to the relation

$$\frac{i\, C_o}{Eh} \propto \frac{r_i^2 \sqrt{2E^*}}{RhE}$$

Thus, if the explosive products control the deformation, the critical deformation criterion implies that

$$\frac{R_o^2}{R} \frac{\sqrt{2E^*}}{h} = constant, \text{ or } \frac{r_i^2 \sqrt{2E^*}}{h} \text{ vs } R \text{ be linear.}$$

(The parameters C_o, ρ, and E, which don't vary in our experiments are suppressed). The data are plotted in Figure 6 and pertinent calculated parameters are reported in Table 1. A regression analysis of R into each of the parameters in Table 1 was made and the correlation coefficients are shown in Table 2. Note that R correlates very well with i/h, but does not correlate significantly with any of the fragment parameters: in interround communication between fragmenting munitions, the development of violent reaction is independent of fragment parameters. That the fragments contribute to the initiation process can be seen by comparing the 50% separation distance for a 105mm HEP acceptor and a bare charge donor of the

Figure 6. The dependence of R upon scaled impulse intensity.

Table I. Measured and Calculated Results for Interround Communication

Donor	Acceptor	R (cm)	E_f (g/s^2)x10^{10} Fragment Areal Kinetic Energy	P_f $(g/m \cdot s)$x10^7 Fragment Areal Momentum	i $(g/m \cdot s)$x10^5 Total Areal Momentum	h (cm) Acceptor Wall Thickness
M1 HE	M1 HE	17.8	5.6	3.6	11.3	1.02
M1 HE	M393A2	25 ± 5	5.6	3.6	11.3	0.4
M393A2	M1 HE	26	5.7	2.4	64.5	1.02
5" 54	5" 54	35.4	3.7	4.1	60.6	1.65
M393A2	M393A2	119.3	5.7	2.4	64.5	0.4

Table II. Correlation Coefficients for Regression of R onto Various Parameters

	R	E_f	P_f	i	E_f/h	P_f/h	i/h
R	1	0.158	-0.548	0.524	0.557	0.227	0.960

same geometry and explosive content. The threshold for the HEP donor
was 119.5 cm, that for the bare charge donor was 8.7 cm. The apparent
ambiguity can be resolved by noting that the initiation of violent
reaction requires casing failure and the explosive to be under compres-
sion when the casing fails. Casing failure is greatly facilitated by
fragment impact, which induces high stresses in the casing, causing
incipient spall. Compression of the explosive is determined by the
deformation of the casing, however, and the deformation is proportional
to the impulse, as shown earlier.

The Vulnerability of Cased Munitions to Impact by Single Fragments

The situation discussed above would change with increasing
fragment mass and velocity; eventually an individual fragment would
have sufficient areal impulse to reach the critical deformation to
cause casing failure. Such is the case for threshold data from gun
firings where there is no loading from explosive products and where the
response of the target is to impact by single fragments. Reeves'
data (6) covered fragment masses from 1.94 gm to 15.55 gm, impacting
against composition B loaded 105mm HE warheads. We sponsored acquisi-
tion of additional data for identical targets, with fragments ranging
from 75 gm to 300 gm. All fragments were steel, right circular
cylinders, with L/D = 1. The response of the targets was inferred
from post-firing examination of a 2.5 cm steel witness plate, and from
recovery of target fragments. The criterion for a violent reaction
was perforation of the witness plate. All data were obtained using

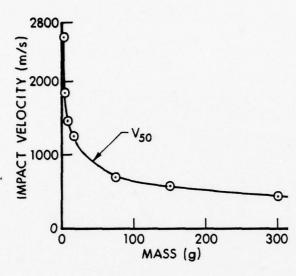a standard quantal response technique (2). The data are plotted in Figure 7, together with Reeves' data. The fact that both sets of data fall upon the same curve is reassuring, and indicates that no experimental artifacts have been introduced because of different lots of munitions and different experimentors.

It is reasonable to expect that the fragment impact initiation of violent reaction in confined targets will obey the same mechanism as does interround communication. If so, the criterion for initiation of violent reaction is that a critical areal impulse/target casing thickness ratio be exceeded. Thus, we have for the 50% threshold locus of mass versus impact velocity, $\frac{mV}{Ah}$ = constant = $\rho \frac{LV}{h}$ where m is the fragment mass, V its impact velocity, A its area, L its length, and h the target casing thickness. Geometric similarily requires that L be proportional to the fragment radius. In particular, for the L/D = 1 fragments used here,

Figure 7. Threshold initiation (50%) for fragment impact of 105mm shell.

$$\frac{mV}{Ah} = \frac{2\rho\ rV}{h}$$

Since

$$r = \frac{m}{2\rho\pi}^{1/3}$$

the criterion becomes

$$\frac{m^{1/3}V}{h} = \text{constant}.$$

The data of Figure 7 are replotted in Figure 8 (circles).
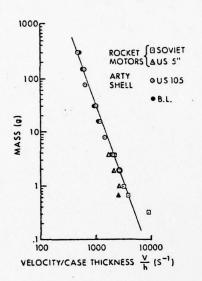
Figure 8. Comparison of initiation thresholds for rockets and shell, ballistic limit of shell.

The solid curve is a straight line with a slope of -3. Note that the fit is good over a three decade change in mass. This provides very strong support for a mechanism which leads to an areal impulse criterion for initiation of violent reaction.

Two corollaries follow from the above results: First, if the condition for initiation of violent reaction is that sufficient deformation of the casing occur to cause failure, i.e., crack generation, then the threshold for initiation should lie very close to the ballistic limit for the casing. To check this, firings against wax filled 105mm shell were conducted using 30 gm, 150 gm, and 300 gm L/D = 1 steel, cylindrical fragments. Some firings of fragments with masses of 3.87 g and 15.45g reported by Reeves (6) are included, also.

The data are shown in Figure 8, where the solid circles represent the average of the highest impact velocity at which no perforation of the casing was obtained and the lowest velocity at which perforation occurred. As can be seen, the conditions for initiation and the ballistic limit are nearly coincident.

The second corollary is that, if the condition for initiation is essentially that casing failure occur, then that condition would apply to different explosives systems, provided that the explosive is not very insensitive (a very insensitive explosive could cause the initiation to depend upon explosive parameters, rather than casing parameters). Some fragment impact data for the US 5" rocket motor and 122mm Soviet rocket motor from (7) are shown in Figure 8. Note that the data are coincident with the 105mm M1 data, in spite of major differences in composition of the filler. The compositions of the rocket motor propellant and the explosive fill are shown in Table 3.

Table III. Chemical Composition

Munition	Composition	Principal Ingredients
105mm M1 HE	B	60% RDX, 40% TNT
105mm M393 HEP	A-3	91% RDX, 9% Wax
5" MK 10 Mod 7 Rocket	double base prop.	51.4% Nitrocellulose, 42.9% Nitroglycerine, 3% Diethylphthalate
122mm Soviet Rocket	double base	(Composition Classified)
Navy 5" 54	A-3	91% RDX, 9% Wax

It is apparent that, at least for those systems for which data exist, the initiation of violent reaction by fragment impact is independent of the filler and is determined by the response of the casing.

Remedial Techniques

An understanding of the mechanism of initiation permits development of techniques which prevent or reduce the frequency of violent reactions resulting from fragment impacts and the detonation of nearby warheads. For munitions in the inventory, protective shields can be developed which reduce the stress levels and stress gradients experienced by the target casings, thus reducing the probability of casing failure. Elementary considerations in shock physics indicate that the best shields are those composed of materials with low shock impedances. The presence of a low shock impedance material between the impacting fragments and the casing causes a more gradual buildup of pressure in the casing and allows more time for rarefactions to reduce the peak stress. Thus, materials such as polyvinyl chloride, foamed metal, etc. should make good shields.

Based on this reasoning, shields were designed to prevent interround communication between 105mm M 456 HEAT warheads contained in a tank ammunition compartment. The effectiveness of the shields relied upon prevention of direct impacts by fragments upon warhead casings and reduction of the shock wave strength experienced by casing and explosive. A series of tests involving two warheads and a single shield per test demonstrated that a 5cm x 5cm x 40cm polyvinyl chloride bar effectively prevented reaction of the acceptors. Thus, in fifteen tests, not a single acceptor warhead detonated, exploded, or showed any evidence of reaction. For these tests, the wall to wall warhead separation was 5 cm. Three further tests were conducted to assess the effectiveness of the shields in a simulated tank ammunition compartment. The compartment geometry, wall thicknesses, interround spacing, etc., closely replicated designs currently under consideration

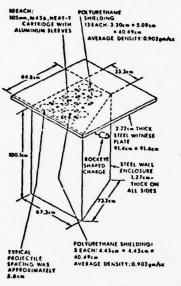for the XM1 tank. A schematic of the setup is shown in Figure 9.
One warhead in each test was deliberately detonated by attacking it with Rockeye shaped charge. In each test, only the deliberately detonated warhead reacted. Results of tests and design information have been provided the project manager, for incorporation into the XM1 tank.

Figure 9. Mockup of tank compartment for confinement effects.

Another, complementary approach can be applied to new warheads entering the inventory. Ignition occurs when the casing fails, and is believed to be caused by the rapid extrusion of the explosive through casing cracks. If this is so, the ignition threshold could be raised by lining the warhead with a thin layer of a pliable polymeric material, which would act as a buffer between explosive and metal interface, during extrusion. (Experiments have shown that ignition occurs more readily as a result of metal- explosive friction than explosive-plastic or explosive-explosive interactions.) To test this hypothesis, 105mm M1 casings were lined with a 3mm coating of cellulose acetate butyrate. The coating thickness was chosen somewhat arbitarily and does not represent a minimum effective thickness. (Note, however, that 3mm is much too thin to provide significant shock attenuation - if shock ignition is the mechanism, the coating will be ineffectual.) Firings were conducted against the polymeric lined munitions with 8 gm steel fragments. The 50% threshold for such a fragment against an unprotected munition is 1470 m/s (4823 f/s). With the lined munitions, no evidence of reaction was obtained at impact velocities of 1740 m/s (5700 f/s), although this is well beyond the ballistic limit of the casing, and perforation occurred. Mild burning reactions were obtained at higher velocities. Even at impact velocities of nearly 2000 m/s (6500 f/s) the warheads did not react with sufficient violence to split open the casings.

SUMMARY AND CONCLUSIONS

Experiments are reported which were conducted to determine the interactions which occur between vicinal munitions. These experiments provided a data base we needed to address mass detonation issues and were designed to provide mechanistic information. Available theory and analysis provided a criterion for initiation, based upon an assumption about the mechanism. The initiation criterion permitted description of the threshold conditions for interround communication. In addition, it was found that, although the fragments participated in interround communication, the process was insensitive to donor fragment parameters, contrary to expectation.

Single fragment impact data was obtained against heavily confined targets. The data base was extended over that available in the literature so that the masses for impacting fragments ranged from 2 to 300 gms. The initiation criterion developed for interround communication was tested against single fragment impact initiation and shown to apply over three decades change in mass, the entire range for which data are available. Both the interround communication data and the single fragment impact data were shown to be consistent with a mechanism which involved deformation of the casing, compression of the explosive, failure of the casing, rapid extrusion of the explosive into cracks, causing ignition, and spread of reaction. Failure of the casing was found to be the critical step, and the initiation criterion was identified with the ballistic limit of the casing.

Since the rupture of the casing controlled the initiation process, the model should be applicable to other systems with similar geometries, but not necessarily similar chemical compositions. This hypothesis was tested against the US 5" MK 10 mod 7 rocket motor and the USSR 122mm rocket motor. Within the accuracy of the data, the initiation criterion applied equally well to these two systems as for the composition B targets for which it was developed.

Additional experiments were conducted to verify the hypothesis that initiation resulted from rupture of the casing and extrusion of the explosive into cracks. Constantan strain gauges and manganin pressure gauges were used to monitor the response of casing and explosive to various stimuli. It was found that catastrophic reaction resulted immediately after casing failure, given a deliberate ignition source. If the samples were not deliberately ignited, but were subjected to rapid deformation, ignition occurred when rupture occurred, with subsequent violent reaction.

HOWE & FREY

Understanding of the mechanism was used to develop remedial techniques. The use of low shock impedance materials to prevent casing fracture was explored and a technique which prevented any inter-round propagation in compartmentalized tank HEAT ammunition was developed. Design information was provided to the XM1 project manager, for incorporation into the new tank.

A technique was developed and tested applicable to munitions entering the inventroy. This technique isolates the explosive from the casing and greatly improves the response of the munition to fragment impact.

REFERENCES

1. Howe, P. et al, "Shock Initiation and the Critical Energy Concept" Sixth Symposium (Int.) on Detonation, San Diego (1976).

2. Dixon, W. J. and Massey, F. J. "Introduction to Statistical Analysis" McGraw-Hill, New York (1969).

3. Liddiard, T "The Initiation of Burning in High Explosives by Shock Waves" Fourth Symp. (Int.) in Detonation, Pasadena, (1965).

4. Taylor, B. BRL (Private communication).

5. Huffington, N. J., Journal of Engineering for Industry, Trans. ASME 1311 (Nov 1975).

6. Santiago, J., BRL R 1571 (1972).

7. Reeves, H., BRLMR 2031 (Mar 1970).

8. BRLCR #65, (Mar 72).

MIRADCOM PROGRAM IN SWEPT-GAIN SUPERRADIANCE (U)

*DAVID W. HOWGATE, DR., CHARLES M. BOWDEN, DR., and JOHN J. EHRLICH,MR.
US ARMY MISSILE RESEARCH AND DEVELOPMENT COMMAND
REDSTONE ARSENAL, ALABAMA 35809

I. INTRODUCTION

Recently, Bonifacio, Hopf, Meystre and Scully (1), using amplifier theory, predicted that under certain conditions, a noise amplifier, pumped by an impulse excitation traveling at the speed of light in the active material, can produce highly nonlinear spacially asymptotic coherent pulses of electromagnetic energy of anomalous intensity and of anomalously short temporal width. The pulse intensities were predicted to increase as the square of the density of the active material, whereas the temporal width should decrease as the inverse of the density. Also, the pulses are characterized by a temporal delay from the pump cutoff to the peak of the pulse evolution. This process has come to be known as swept-gain superradiance.

Electromagnetic energy having these general characteristics would be useful in application to LADAR systems for propagation, ranging, discrimination and coherent imaging. Other applications would include coupling to plasmas for efficient energy delivery needed for laser induced fusion and for plasma diagnostics. Since the pulses are produced in the amplifier configuration without mirrors, this presents an attractive scheme for production of unidirectional coherent VUV and X-ray radiation.

Because of the strong potential application to Department of the Army missions, MIRADCOM scientists first improved upon the theoretical model of Bonifacio et al. (1) using coherent pumping on molecular species for realistic conditions dealing explicitly with pump pulse characteristics, temporal width and propagation (2). This

1

was done for the purpose of experimental design and interpretation of results.
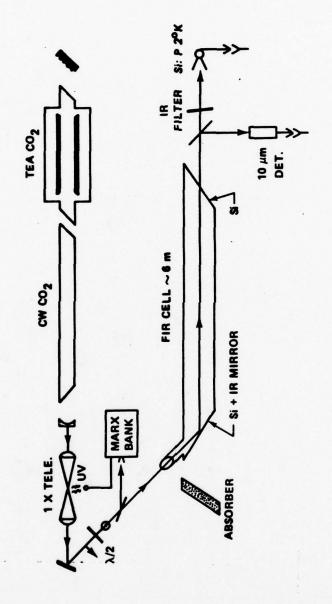
Using the results of our theoretical model, an experiment was then designed and performed at MIRADCOM using CO_2 pumped CH_3F. In this effort MIRADCOM scientists collaborated with University of Illinois scientists to build the apparatus and to perform the experiments at MIRADCOM. The initial results of this work were reported at an international meeting and will appear in the publication of the Proceedings (3). An updated theoretical development and experimental results was also delivered by one of the authors as an invited paper at the 8th Annual Colloquium on the Physics of Quantum Electronics, Snowbird, Utah, January 15-17, 1978. The final version of our recent work will be published shortly (2,4).

In the next section, the experiment will be described and the principal results given. Section three will be devoted to a discussion of the essentials of the theoretical model and interpretation of the experimental results. The last section will be used to present conclusions drawn from our initial work and to point out the future direction of effort. Further details of the experiment are given in references (3,4).

II. THE MIRADCOM EXPERIMENT
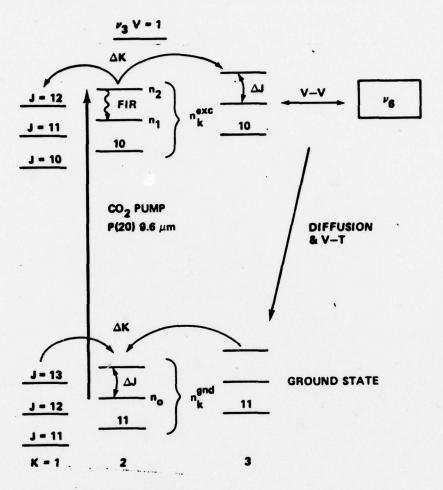
The experimental arrangement for the MIRADCOM experiment is based upon results of our theoretical model and the earlier experimental studies of Dicke superradiance in CH_3F by Rosenberger, Petuchowski and DeTemple (5). The schematic for the experiment is presented in Figure 1.

The TEA CO_2 laser is operated on the P(20) ($\lambda = 9.55$ μm) line to give a single longitudinal and transverse mode by using a low pressure CW CO_2 gain cell. Smooth 200 nsec pulses are produced of about 150 mJ total energy. These pulses are subsequently chopped using an optical breakdown switch which utilizes UV-triggered AC breakdown of clean N_2. The resulting CO_2 pulse is about 65 nsec duration and is cut off on its trailing side in less than 0.1 nsec. This pulse is passed through a CH_3F cell, 6 m in length along the cell axis. The area of the IR beam is approximately 2 cm^2. The resulting FIR pulses at 496 μm were monitored in both the forward and backward directions by a low temperature phosphorous doped silicon detector which operated near 2°K. At the output end of the CH_3F cell the IR pump pulse which emerges and the FIR pulse generated were monitored simultaneously by using a beam splitter which reflects nearly 100% of the IR and

(Printed by permission of
Plenum Press)

Figure 1. Experimental Arrangement

(Printed by permission of
Plenum Press)

Figure 2. CH_3F Energy Levels Relevant to
the Experiment

4

transmits about 50% of the FIR. These pulses were displayed in the same time frame on an oscilloscope, whereby the relative temporal positions of the two pulses were easily determined. Further details of the experimental apparatus are given in references (3) and (5). The relevant energy level diagram for CH_3F for the transition involved in the experiment is presented in Figure 2. Since it was possible to couple to two k-level transitions (approximately 45 MHz apart) by tuning the pump cavity, we tuned to the one corresponding to the highest gain, i.e., k = 2.

Results of this experiment were reported earlier (3) and are presented in Figures 3, 4 and 5. In Figure 3(a), the FIR forward pulse intensity at 496 μm is plotted vs the square of the pressure in the CH_3F cell and confirms the linear dependence predicted from theory. In this figure and all subsequent representations of the data, each data point represents the average of three scope traces taken in succession. In this pressure range (P < 0.2 Torr), the forward and backward FIR wave intensities are equal. Shown in Figure 3(b) is the FIR forward intensity vs the square of the pressure for the data of Figure 3(a) on a different scale, and including measurements at higher pressure (P > 0.2 Torr). The break in the pressure dependency of the intensity occurs at approximately 0.2 Torr. Also represented in Figure 3(b) is the ratio of forward to backward wave intensity (dashed curve and right-hand scale) obtained from the curve of Figure 3(c) through the data points. It is seen from Figure 3(b) that the forward wave intensity becomes greater than that of the backward wave for pressures greater than 0.2 Torr, the point at which the break occurs in the slope of the forward intensity vs P^2 curve. Thus, for P > 0.2 Torr the emitted pulses are generated by swept-gain superradiance (1), whereas for lower pressure, P < 0.2 Torr (where the forward to backward wave intensity ratios are equal), the pulses are generated by Dicke superradiance (6). The latter has been studied and reported earlier by Rosenberger, Petuchowski and DeTemple (5).

The temporal width of the pulses as a function of inverse pressure is shown in Figure 4(a). The break occurs at P ≈ 0.09 Torr, which marks the transition from inhomogeneously broadened (lower pressure) to homogeneously broadened (higher pressure) regimes. The pulse width is seen to decrease with increasing pressure and it actually becomes smaller than the absorption line width in the higher pressure homogeneously broadened regime.

Shown in Figure 4(b) is the pulse delay, measured from the IR pump cutoff to the FIR pulse peak intensity, plotted vs inverse pressure. The FIR pulse is well separated from the IR pump (temporally) at lower pressures and narrows in width, grows
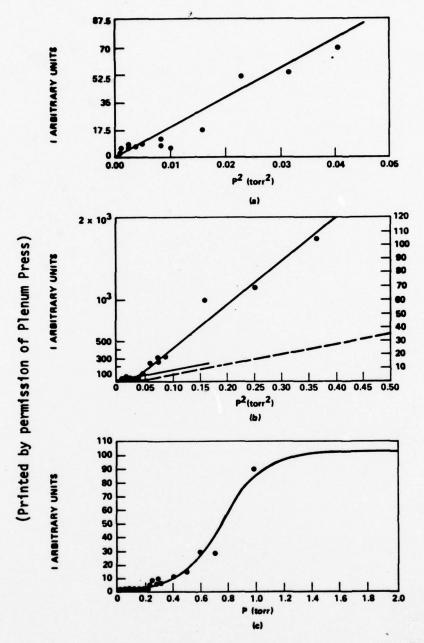
(Printed by permission of Plenum Press)

Figure 3. FIR Intensity I as a Function of Pressure P .

6
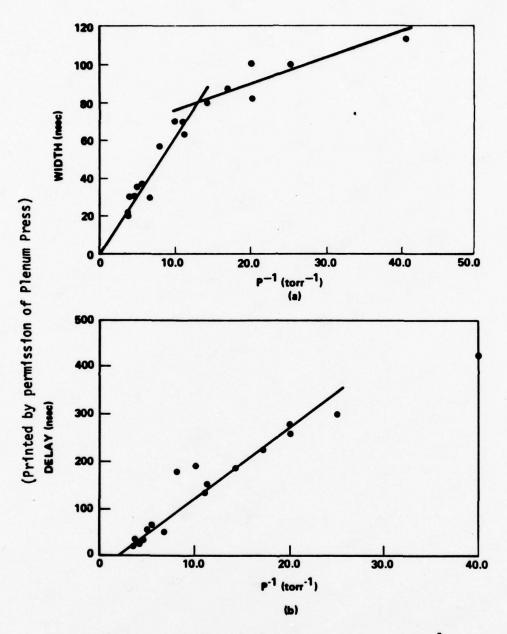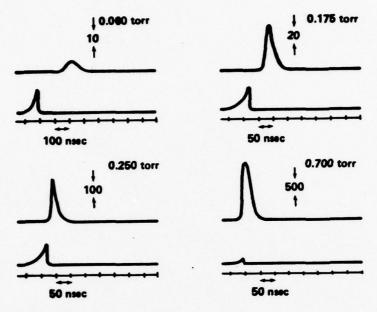
Figure 4. FIR Width and Delay vs Inverse Pressure P^{-1} .

Figure 5. FIR (Upper) and IR (Lower) Scope Traces

nonlinearly in intensity and moves closer to the pump pulse as the pressure is increased. The negative intercept is due to the fact that the IR pump pulse is not an impulse but is a coherent pulse of non-zero duration, giving rise to an FIR pulse evolution which begins inside the pumping pulse.
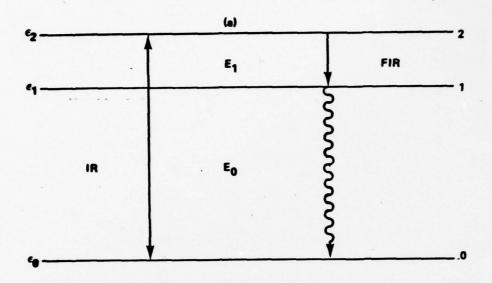
Actual scope traces of the IR pump and superradiant FIR pulses are presented in Figure 5. The pulses are presented in the same time frame for four different CH_3F cell pressures. The upper pulse is the FIR at 496 μm and the lower one is the IR pump pulse at 9.6 μm as received by the respective detectors at the forward end of the CH_3F pressure cell (as shown in Figure 1) and subsequently displayed simultaneously on the oscilloscope in the same time frame. The pulses are both traveling to the left in the figure and the vertical scale for the IR pump (lower trace) is the same throughout. The FIR superradiant pulse is seen to become narrower and nonlinearly more intense and move in toward the pump pulse as the pressure

8
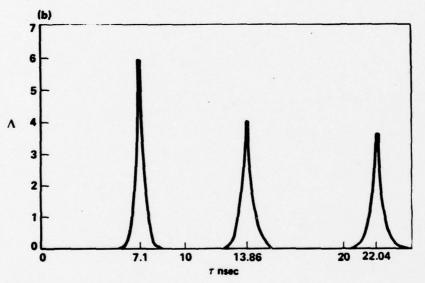
becomes greater. At the highest pressure shown, the lower right-hand traces, the FIR and IR temporally overlap. In this regime, the FIR superradiant pulse continues to increase in intensity as the square of the pressure even though the pump energy is markedly depleted. For the case shown, there is less than 10% of the pump pulse energy remaining. The FIR pulse intensity continues to grow nonlinearly until virtually all of the pump energy is gone. It is apparent that in this case the maximum conversion of IR energy to FIR pulse energy occurs consistent with conservation of energy.

III. THEORETICAL MODEL AND INTERPRETATION OF THE EXPERIMENT

The theoretical model is based upon the energy level diagram shown in Figure 6(a). The IR pump is treated as an external coherent source of prescribed shape and temporal duration, propagating at the speed of light in the medium and transfering population from the ground state to the excited state. The subsequent transition of population from the excited state to the intermediate level gives rise to the FIR radiation which evolves. The FIR radiation field is treated quantum mechanically to incorporate spontaneous as well as stimulated decay. Although the intermediate and ground states are not radiatively coupled, because of rotational selection rules in the molecule, (resonant) stimulated Raman transitions transfer population between them and have important manifestations in the results in terms of the time evolution of the gain. It should be noted that these results for coherent pumping pertain to general laser theory as well and have not (to our knowledge), been treated in the open literature. The model incorporates the aspects of multimode dependence for the FIR pulse evolution as well as propagational dependence of both the pump IR and FIR pulses. Due to limitations in space, only one aspect of the results of the model will be presented here. The model and the detailed comparison with the experiment are discussed fully in the references (2,9).

The calculated behavior of the FIR pulse evolution in the high pressure regime when pump pulse and FIR pulse temporally overlap is presented in Figure 6(b) for a set of realistic system parameters. The pump is turned on at $\tau = 0$ (coordinate system moving with leading edge of the pump) and remains on at constant amplitude throughout the plot. What is shown is the evolution of the FIR pulse from $\tau = 0$, where the vertical axis is the FIR intensity in units of IR pump intensity. This predicts that if the pump pulse is sufficiently long, a set of sharp, intense pulses will be generated in the FIR which are phase correlated from one pulse to the next. The damping in amplitude is due to the dephasing time of the transition. This also shows

9

(Printed by permission of
Plenum Press)

Figure 6(a) Model Energy Level Scheme; (b) Calculated FIR Pulse
Sequence Evolution vs Time τ from IR Pump Turn-on for a Step
Function Pump Envelope.

that one can achieve short (2 nsec) pulses of greater intensity than the pump pulse. This does not violate conservation of total energy since the pulses generated in this manner are "off" more than they are "on" during the duration of the pumping pulse.

Pulse chains of this type have actually been observed (3) in our experiments when we have temporally extended the pump by turning off the N_2 gas breakdown switch (which chops the pump pulse in the experiment, Figure 1). Pulse chains of this type should be quite useful in many field applications where coherent, short and intense pulses are desired.

IV. CONCLUSIONS

MIRADCOM scientists have demonstrated for the first time swept-gain superradiance. It was shown experimentally that pulses of anomalous intensity and short temporal widths could be produced and the width and intensities controlled by varying the pressure in the cell. Furthermore, it was predicted by theory and experimentally observed (3) that sets or chains of phase coherent pulses of the superradiant type can be generated by coherent pumping. Pulses of this type have obvious Department of the Army mission applications.

Further research is planned for continuing development of the theory (for deeper understanding of the phenomenon) and further experiments are planned to confirm predictions and to determine practical limitations on power density, pulse compression and pulse shaping. A portion of this research is presently underway. Peak powers of 1-2 megawatts and pulse durations on the order of nanoseconds are not unreasonable to anticipate at submillimeter wavelengths.

It is also our intention to carry these studies to materials giving rise to shorter wavelength radiation consistent with the myriad of applications which are possible. We intend to determine the effects of incoherent pumping (Bluemline discharge (7)) on swept-gain superradiance pulse evolution for optical and VUV wavelength pulse generation.

Our initial work opens a vast area of investigation over a broad range of wavelengths for the production of coherent, intense, short pulses generated from noise (and without mirrors) which can be shaped and compressed by control of the pressure and/or cell length. Theory predicts that pulses of this nature will be produced in any medium for which the gain to loss ratio is greater than unity and for which the total gain is greater than one.

*HOWGATE, BOWDEN and EHRLICH

ACKNOWLEDGEMENTS

REFERENCES

1. R. Bonifacio, F. A. Hopf, P. Meystre and M. O. Scully, Phys. Rev. A$\underline{12}$, 2568 (1975).

2. C. M. Bowden and C. C. Sung, to be published; C. M. Bowden, C. A. Coulter, and D. W. Howgate, to be published.

3. J. J. Ehrlich, C. M. Bowden, D. W. Howgate, S. H. Lehnigk, A. T. Rosenberger, and T. A. DeTemple, "Swept-Gain Superradiance in CO_2-Pumped CH_3F", Proceedings of the Fourth Rochester Conference on Quantum Optics, ed., L. Mandel and E. Wolf, Plenum Press, 1978, in press.

4. C. M. Bowden, J. J. Ehrlich, T. A. DeTemple, and A. T. Rosenberger, to be published.

5. A. T. Rosenberger, S. J. Petuchowski and T. A. DeTemple, "Experiments in FIR Superradiance" in Cooperative Effects in Matter and Radiation, edited by C. M. Bowden, D. W. Howgate and H. R. Robl, Plenum Press, New York, 1977, pp. 15-35.

6. R. H. Dicke, Phys. Rev. $\underline{93}$, 99 (1954). See also for criteria: J. C. McGillivray and M. J. Feld, "Superradiance in Experimentally Relevant Regimes" and R. Bonifacio, M. Gronchi, L. A. Lugiato and A. M. Ricca, "Maxwell-Bloch Equations and Mean-Field Theory for Superfluorescence" in Cooperative Effects in Matter and Radiation, edited by C. M. Bowden, D. W. Howgate, and H. R. Robl, Plenum Press, New York, 1977, pp. 1-15; 193-209.

7. J. D. Shipman, Appl. Phys. Lett. $\underline{10}$, 3 (1967); J. N. Knyazer, V. S. Letokhov and V. G. Marshev, IEEE Jour. Quantum Elec. $\underline{QE-11}$, 805 (1975).

A COATING AGENT FOR PROMOTION OF METAL-TO-PORCELAIN BONDING

EUGENE F. HUGET, LTC, DC, MR. LASZLO B. DE SIMON
UNITED STATES ARMY INSTITUTE OF DENTAL RESEARCH
WASHINGTON, D. C. 20012

The fusion of porcelains to metals is an ancient art. Although its specific origin cannot be substantiated, it is known that metal ceramic decorative and functional devices were fabricated in many early civilizations, particularly those of the Egyptians, Greeks and Persians.

Since the early 1950's, esthetic porcelain-fused-to-metal dental treatment devices have enjoyed wide clinical use. The following problems, however, initially encumbered the production of porcelain-metal dental restorations: (1) Fused porcelain displayed cracks after firing because of the mismatch of thermal expansion between the porcelain and the metal substructure; (2) the color of fused porcelain was altered adversely by color forming oxides and by technique; (3) relatively soft gold-based casting alloys lacked sufficient strength and rigidity to withstand rigorous functional service.

Over the past decade, the foregoing difficulties have been reduced in frequency of occurrence and in severity (1,2). Today, the degree of bonding of fused porcelains to improved gold-containing (precious) alloys appears to be adequate. However, the mechanisms by which bonding occurs are not understood completely. Several explanations of the so-called porcelain-to-metal bond have been advanced. Van der Waal's forces (3,4), and the beneficial effects of compression forces resulting from small differences in thermal expansion of two dissimilar phases may contribute to bond formation.

In recent years, costs incurred in the laboratory fabrication of porcelain-fused-to-metal fixed prosthetic devices (crowns and bridges) have increased markedly. Increased production costs have been, for the most part, reflections of the advancing international

market values of gold and other precious metals. Searches for less
expensive veneerable metallic substrates from which functional as well
as esthetic restorations might be fabricated have led to interest in
the use of materials alloyed from base metals rather than from pre-
cious constituents. Nickel (~60 to 80 percent) and chromium (~12 to
20 percent) are major components of the majority of available substi-
tutes for high-fusing gold-based dental casting alloys. Although the
physical and mechanical properties of the nickel-chromium alloys are
sufficient for functional intraoral usage, difficulties have been ex-
perienced in obtaining reliable porcelain-to-metal bonding (6).

It would appear, from laboratory and clinical observations,
that a relatively thick scale which accrues at the substrate surface
during the application of the esthetic ceramic veneer is not conducive
to the formation of color-stable and strong porcelain-to-metal bonds.
Pretreatment of metallic substructures by the application of colloidal
gold or by metal-ceramic mixtures prior to the application of porce-
lain presents an intriguingly expedient means for achieving improved
bonding of fused dental porcelain to nonprecious (nickel-chromium)
casting alloys. However, the excessive expense of colloidal gold
preparations and the highly critical parameters of technique demanded
for the effective use of proprietary metal-ceramic mixtures have limi-
ted the scope of acceptance of adhesion-promoting coating agents for
base metals.

Tasks directed toward the development of inexpensive coating
agents and the refinement of techniques for the promotion of nonpre-
cious metal-to-porcelain bonding are being conducted at the United
States Army Institute of Dental Research. This report provides an
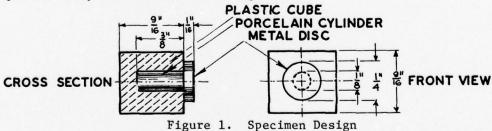up-to-date summary of our progress.

MATERIALS AND METHODS

A simple test for rapid measurement of the apparent bond
strength of metal-porcelain couples was developed within this labora-
tory. The metallic components of the test pieces were 1/4 X 1/16-inch
discs. The discs were cast from six proprietary base metal alloys*
in accordance with the laboratory procedures recommended by the re-
spective manufacturers of the materials. Two series of castings for
each of the alloys were used in specimen preparation. Test surfaces
of the member castings of one series were handground on 240-grit
metallographic papers. Then the discs were subjected to a 5-minute

___
* Neydium, J. M. Ney Co., Hartford, CN 06101; Gemini II, Kerr Sybron
Corp., Romulus, MI 48174; NP-2, Howmedica, Inc., Dental Division, Chi-
cago, IL 60632; Omega-VK, Pro-Met, Inc., Oak Brook, IL 60521; Ticon,
Ticonium Co., Inc., Albany, NY 12207; Ceramalloy, Johnson & Johnson
Dental Products Division, East Windsor, NJ 08520.

heat treatment at 1,950° F., cooled to room temperature in open air and again ground lightly on 240-grit abrasive paper. Castings included in a subsequent series were also surface-finished with 240-grit papers. Test surfaces of these discs were covered with a thin coating produced from a slurry of ethyl alcohol and a metal-ceramic powder. Composition of the powder component of the coating material was fine aluminum (1 part by weight) and opaque dental porcelain (5 parts by weight). The coated discs were heated to 1,900° F. and cooled to room temperature. Sintered remnants of the metal-ceramic coating were scraped from the discs to create relatively flat surfaces on which porcelain could be condensed.

Segments of 1/8-inch inside diameter polyethylene tubing approximately 3/16 of an inch in length were attached to the castings with the use of rubber cement. This arrangement provided adequate matricies into which aqueous slurries of the opaque powder component of a dental porcelain[+] could be placed. After thorough condensation of the porcelain, the specimens were set aside and allowed to dry in room air for 20 minutes. Then the assemblies were placed in an open muffle for ignition and thermal destruction of the tubing.

The specimens were placed inside the muffle after temperature stabilization at 1,200° F. was achieved. The furnace temperature was increased at a rate of 80° F. per minute until the porcelain-fusion temperature of 1,825° F. was reached. The muffle vacuum was broken and the specimens were held at 1,825° F. for two minutes. The fused components were removed from the muffle and cooled to room temperature in open air. The fused alloy-porcelain pieces were placed in 9/16-inch cubic silicone rubber molds. Cylindrical receptacles at the bases of the molds insured the desired orientation of the alloy-ceramic specimens. The mold cavities were filled with a commercial dental tray acrylic.[#] A schematic representation of the fused and mounted specimen-components is shown in Figure 1.

Figure 1. Specimen Design

+ B. F. Vacuum Porcelain (Paint-O-Pake), Ceramco, Inc., New York, NY 11101.
# Fastray, Harry J. Bosworth Co., Chicago, IL 60605.

HUGET, DE SIMON

Each cube was fastened in a small vise for added support. The assembly was oriented on a testing machine[§] to allow loading of the cast disc in a diametral direction. Crosshead speed of the testing machine was 0.02 inch per minute. Bond strength was calculated on the basis of ultimate load at shear failure per unit area of apparent porcelain-to-metal contact.

RESULTS

Load tracings (Figure 2) obtained on fracture separation of the metal-porcelain couples did not show deformation prior to bond failure.
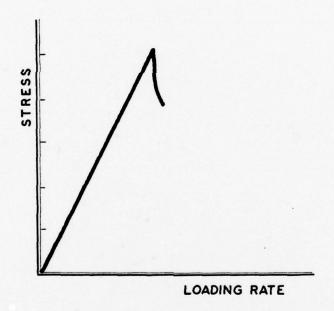
Figure 2. Typical load tracing.

§ Instron Universal Tensile Testing Machine, Instron Corp., Canton, MA 02021.

A prominent feature of the pattern of breakage displayed by specimens made from noncoated discs is shown in Figure 3. A black scale remained attached to the contact surface of the porcelain cylinder. It appeared that breakage occurred between the metal substrate and an oxide layer on its surface. All specimens produced from noncoated discs of the six test alloys failed in this manner when subjected to shear loading.

Failure of specimens made from aluminum-ceramic coated castings occurred in the vicinity of the coating-porcelain phase boundary (Figure 4). Remnants of porcelain of varying thickness remained attached to the coated surfaces of the metal discs.

Bond strength data for specimens fabricated from noncoated and coated cast discs are presented in Tables I and II, respectively. All reported values are based on a minimum of 30 observations.

Specimens made with the use of noncoated castings exhibited relatively poor porcelain-to-metal bonding. The six nickel-chromium based alloys gave bond strengths that ranged from 2,700 to 6,500 psi. Precision of the bond strength measurements was low. Coefficients of variation ranged from 29 to 42 percent.

A marked increase in measurement precision was obtained upon coating of the castings prior to the application of porcelain. Coefficients of variation declined to a narrow range of 11 to 17 percent. Also, specimens fabricated from coated castings gave significantly higher bond strength values than specimens made from noncoated castings. Increases in the apparent strengths of the metal-porcelain systems studied ranged from 10 percent to 66 percent for Gemini II-porcelain and Neydium-porcelain combinations, respectively.

DISCUSSION

The application of porcelain to base metal alloys is technique sensitive. Failure of the porcelain-metal bond often occurs during the fabrication of a dental restoration. Such a failure is depicted in Figure 5. Separation of the porcelain from the underlying cast substructure is evidenced by the crevice at the occlusal juncture of the metal and the ceramic veneer. The crevice developed while the restoration was being cooled to room temperature following fusion of the porcelain.

Porcelain-fused-to-metal restorations made with the use of the aluminum-ceramic coating seldom exhibit separation defects (Figure 6).

The coating has been employed successfully in the fusion of porcelain to a wide variety of base metal dental treatment devices.

Aqueous slurries of dental porcelain powders can be applied to coated castings with the use of common condensation techniques. Fusion of opaque, body, incisal and gingival porcelains as well as the final glaze can be accomplished without deviation from the time-temperature sequence prescribed by the manufacturer of the selected ceramic materials.

Improved porcelain-to-metal bonding that can be attained through the use of an aluminum-ceramic coating broadens the range of application of inexpensive nickel-chromium alloys in military dental practice.

SUMMARY AND CONCLUSIONS

A coating for enhancement of alloy-porcelain bond strength has been developed and its feasibility demonstrated. The development provides a composition of matter comprising substantially one part by weight fine powdered aluminum, substantially 5 parts by weight opaque dental porcelain, and sufficient water to form a useable aqueous slurry. The composition may include ethyl alcohol and any other liquid which will promote wetting of the substrate surface by the aluminum-ceramic admixture.

Improved base metal to porcelain bonding that can be attained through the use of the aluminum-ceramic coating material broadens the range of application of inexpensive nickel-chromium alloys in military dental practice. The advantages of this coating composition over prior art products are its low cost, ease of preparation and manipulation, and effectiveness in promotion of metal-porcelain bonding.

REFERENCES

1. Howell, R. A.: Gold-Porcelain Bridge-Work. Brit. D. J. 116:80, 1964.

2. McCormack, J.: Porcelain Bonded to Precious Metal Castings. Brit. D. J. 116:213, 1964.

3. O'Brien, W. J., and Ryge, G.: Contact Angles of Drops of Enamel on Metal. J. Prosthet. Dent. 15:1094, 1965.

4. Knap, F. J., and Ryge, G.: Study of Bond Strength of Dental Porcelain Fused to Metals. J. Dent. Res. 45:1047, 1966.

5. Vickery, R. C., and Badinelli, L. A.: Nature of Attachment Forces in Porcelain-Gold Systems. J. Dent. Res. 47:683, 1968.

6. Huget, E. F., Dvivedi, N. N., and Cosner, H. E., Jr.: Properties of Two Nickel-Chromium Crown-and-Bridge Alloys for Porcelain Veneering. J.A.D.A. 94:87, 1977.
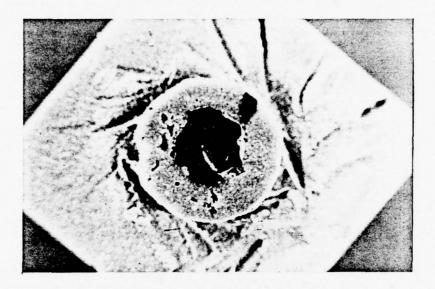
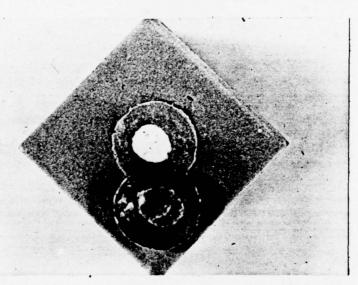Figure 3.   Breakage pattern of specimens
(noncoated discs).

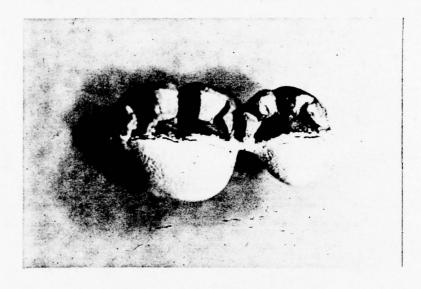
Figure 4.   Breakage pattern of specimens
(coated discs).

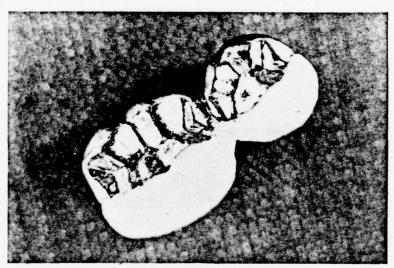7

Figure 5. Metal-porcelain bond failure.

Figure 6. Esthetic fixed prosthetic device. Base metal sub-
structure coated with aluminum-ceramic mixture
prior to application of porcelain veneer.

TABLE I

APPARENT ALLOY-PORCELAIN BOND STRENGTH
NONCOATED CASTINGS

Alloy	Mean Bond Strength (PSI)	C.V. %
Neydium	2,700	34
Gemini II	6,500	29
NP-2	5,600	22
Omega-VK	5,500	33
Ticon	3,200	42
Ceramalloy	4,800	30

TABLE II

APPARENT ALLOY-PORCELAIN BOND STRENGTH
COATED CASTINGS

Alloy	Mean Bond Strength (PSI)	C.V. %
Neydium	7,800	12
Gemini II	7,200	17
NP-2	7,700	14
Omega-VK	7,100	11
Ticon	7,700	16
Ceramalloy	8,400	11

TUNNEL DESTRUCTION STATE-OF-THE-ART

CHARLES E. JOACHIM
U. S. ARMY ENGINEER WATERWAYS EXPERIMENT STATION
VICKSBURG, MISSISSIPPI   39180

Current Army tunnel destruction criteria (1) are based on limited nuclear-explosive (NE) and high-explosive (HE) test data. The PILE DRIVER (2) and HARDHAT (3) events investigated tunnel and support system response where yields were relatively large and tunnel diameter small compared to typical underground openings and localized stress fields of low yield weapons. Some limited tunnel destruction information was obtained from underground events of the PLUMBBOB and HARDTACK II Series (4). These events used access drifts which were designed to be self closing at the weapon point.

The limited NE data are augmented by HE results from the Underground Explosion Test (UET) Program, a series of model and prototype-scale experiments in granite and sandstone (5). Weapon standoff distances (charge c.g. to tunnel wall) were at or near the maximum for major damage. Limited tunnel destruction was produced by a series of hasty and deliberate tests (6) on abandoned railroad tunnels in basalt. Additional HE data were obtained from model tests at the Waterways Experiment Station (WES) (7) using 2-pound TNT charges and tunnel diameters varying from 1.2 to 6.6 inches. Recent laboratory scale HE experiments (8) conducted at the WES under the ESSEX program investigated the effects of tunnel diameter, standoff distance, charge confinement, material strength and to a limited extent, tunnel spacing and liner strength variations on tunnel damage.

This paper summarizes the analysis conducted to develop predictions for the effects of low-yield nuclear weapons against tunnels and underground openings in rock.

DISCUSSION

During the data analysis phase, Hopkinson's (cube root) yield scaling and NE-HE equivalence (for fully contained detonations) were assumed. This analysis showed good correlation between the UET (HE) and HARDHAT (NE) tunnel damage data. These detonations were at or near the maximum standoff distance for major damage. The remaining NE damage data exhibit considerable scatter. These NE events were conducted at the Nevada Test Site (NTS) in tuff, a weakly cemented highly variable material. The data scatter is probably due to local variations in material properties or water contents. HE damage data from detonations inside tunnels (basalt) and vertical shafts (sandstone) demonstrate that this weapon placement option produces only superficial tunnel damage unless very large yields are employed.

The current Army tunnel damage classification system (1) and typical damage profile are depicted in Figure 1. The four damage zones (1, 2, 3, and 4) in this system are the zone of complete damage, the zone of rock breakage, the zone of continuous slabbing and the zone of discontinuous damage, respectively. Damage to the outer limit of Zone 2 is classified as severe and constitutes closure in this system.

Tunnel closure data from medium strength models (8) (tunnel diameter 50 feet*) are plotted versus standoff distance in Figure 2. Symbols denote charge depth of burst (DOB) groupings. Lines are shown which connect maximum closure data for each group, forming envelopes of maximum closure length. These data indicate a general trend of increased length of tunnel closure with increased DOB. Charge tangent to (and above) the model surface (negative DOB) did not produce closure. Shallow charge DOB's (14 to 15 feet) resulted in significant tunnel closure. Comparison between the shallow and intermediate DOB (38 to 130 feet) data shows that the intermediate detonations resulted in a 25 percent increase in the length of tunnel closed. Comparison between the intermediate and deeper DOB (160 to 180 feet) data indicates that the deep detonations produced no significant increase in tunnel closure lengths within the data scatter.

Figure 2 also shows that for medium strength material, very little damage occurred at standoff distances greater than 130 feet and that the optimum standoff distance for producing tunnel closure in this material is in the range of approximately 40 to 100 feet. At smaller standoff distances there was not enough material blown into the tunnel to produce closure or an appreciable obstacle.

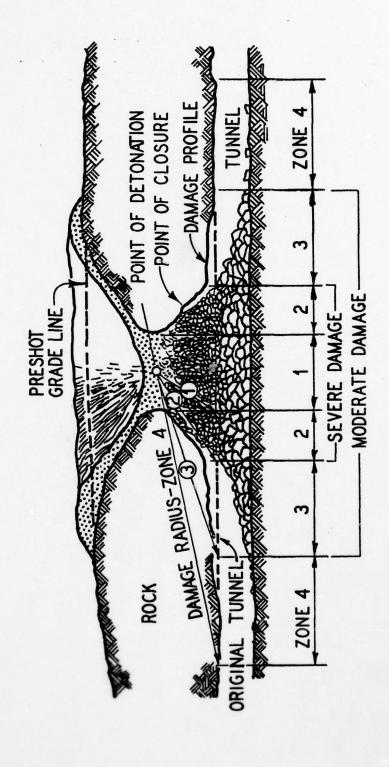*All dimensions given in this discussion are scaled to 1 kiloton.

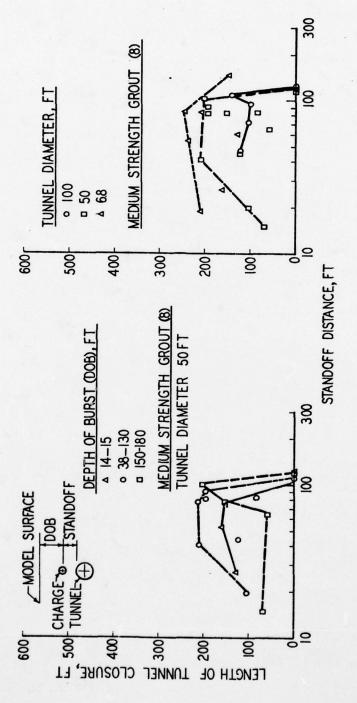Figure 1. Typical tunnel damage profile and damage zones (1)

Figure 3. Effects of tunnel size on damage scaled to 1 KT

Figure 2. Effect of depth of burst on tunnel closure scaled to 1 KT

The effect of tunnel size on length of tunnel closure is presented in Figure 3. Data for three tunnel diameters (100, 59, and 6.8 feet are shown. The maximum standoff distance at which closure occurred was essentially independent of tunnel diameter. As standoff distance decreased a minimum was reached for the larger tunnels where the debris volume was insufficient to fill the tunnel. (Although there was considerable data scatter, the tunnel closure remained relatively constant over a range of standoff distances from approximately 40 to 100 feet.) There was a distinct trend for more damage, as indicated by closure length, as tunnel size decreased.

Tunnel closure data for low, medium and high strength models for tunnel diameters in the range of 50 feet are plotted versus standoff distance in Figure 4. Also shown for comparison are the UET (granite and sandstone), HARDHAT (granite) and NTS (tuff) data. As shown here, these data are in good agreement with the low strength model results, although there is considerable scatter in the NTS (tuff) results.

The maximum standoff distance at which closure occurred increased with decreasing material strength. As shown in Figure 4, six times the length of tunnel was closed in the low strength as in the high strength material for similar test geometries.

A comparison between lined and unlined tunnels is presented in Figure 5. These tests were conducted with the medium strength material and one tunnel diameter (50 feet). Aluminum tubing cast into the model served as tunnel liners for these experiments. Three thicknesses of tubing were used giving thickness to inside diameter (t/d) ratios of 0.0090, 0.029, 0.033, and 0.060. As shown here, tunnel liners reduced the maximum standoff distance at which closure occurred and the length of tunnel closed. These meager data also indicate a tendency for decreased damage with increased liner thickness. The thinnest liner (t/d of 0.0090) failed by buckling over an appreciable portion of its length but did not close completely.

Five tests were conducted in the medium strength material to study the effects of parallel tunnel spacing on damage. Tunnel diameter (50 feet) and charge standoff distance (63 feet) were held constant for these experiments. Tunnel closure length from these experiments is plotted versus the ratio of tunnel spacing (center to center distance) to tunnel diameter (s/d) in Figure 6. A significant increase (approximately 50 percent) in the length of tunnel closed occurred for s/d ratios slightly greater than unity. As the s/d ratio approached 2, the length of tunnel closed fell within the data
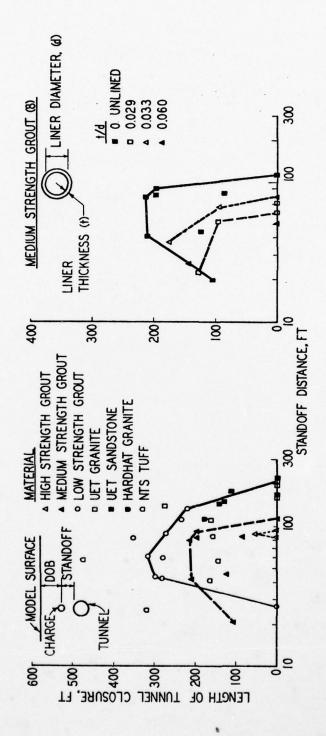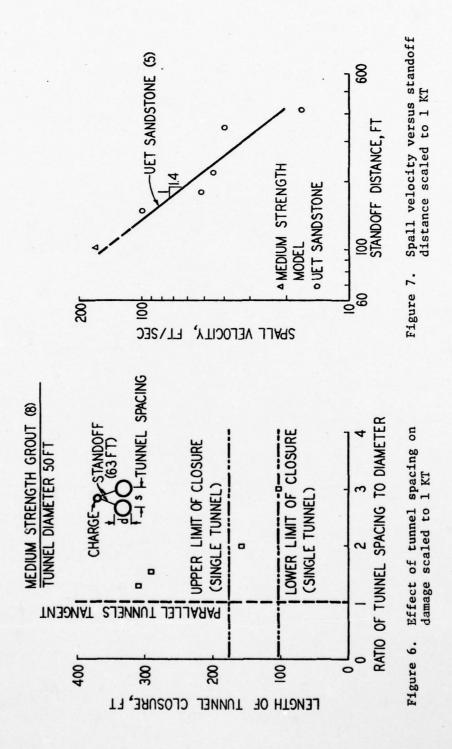
Figure 4. Effect of material strength on damage scaled to 1 KT

Figure 5. Influence of liner thickness on damage scaled to 1 KT

Figure 7. Spall velocity versus standoff distance scaled to 1 KT

Figure 6. Effect of tunnel spacing on damage scaled to 1 KT

scatter experienced from single tunnel at this standoff. The largest
s/d ratio possible with this tunnel diameter-standoff distance
combination is 3.5.

A spall velocity measurement made in a model test is compared to
the UET sandstone results in Figure 7. The sandstone curve is ex-
tended (dashed line) into the region of the model. The spall velocity
data point is the peak value of an integrated acceleration time
history. Good correlation is seen between the model and the UET spall
velocity data.
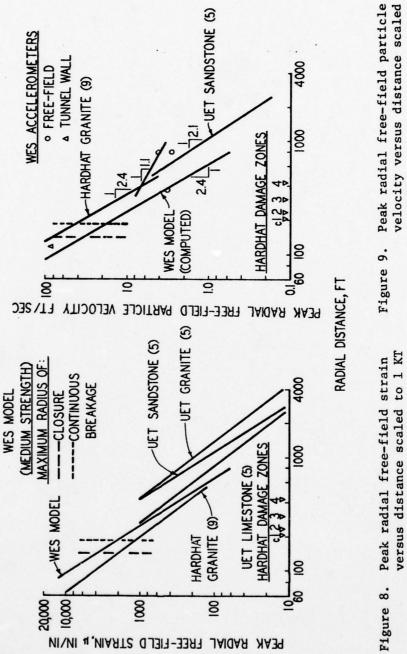
A comparison of least-square fits to the WES model (medium
strength material), UET (limestone, granite and sandstone), and
HARDHAT (granite) radial free-field strain data is shown in Figure 8.
These strain curves are indicative of the tunnel input loading (times
the appropriate concentration factor) which produced the varying
degrees of tunnel damage. Damage zone limits are also shown for the
WES model (dashed vertical lines) and HARDHAT (triangles). Calculated
(from data fit) peak free-field radial strains in the WES model were
3850 and 2050 micro-inches/inch for the maximum radius for closure and
continuous breakage, respectively.

Peak radial particle velocity data from the free-field test block
are presented in Figure 9. The model tunnel peak particle velocity
data point shown here was calculated assuming a free-field velocity
of one-half the spall velocity. Assuming the one-dimensional relation
$v = \varepsilon c$ where c is 9330 ft/sec, the WES model peak free-field radial
strain curve from Figure 8 was used to calculate the velocity curve
shown in Figure 9. Also included in this figure are data fits from
the larger UET events (rounds 814, 815, 816, and 817) in sandstone and
the HARDHAT experiment. Although a factor of two difference exists
between the WES model calculated velocity and the HARDHAT curve,
closure for both experiments occurred at a peak free-field particle
velocity of approximately 40 ft/sec.

CONCLUSIONS

Tunnel damage, as indicated by the degree of closure, is highly
dependent on the strength of the rock; significantly greater damage
is associated with the weaker materials (Figure 4).

Within the bounds of the test conditions, the smaller tunnels
underwent the greatest damage for a given weapon standoff distance
(Figure 3).

Figure 8. Peak radial free-field strain versus distance scaled to 1 KT

Figure 9. Peak radial free-field particle velocity versus distance scaled to 1 KT

For a 1 KT weapon and a tunnel diameter of 50 feet the optimum standoff distance for tunnel closure is in the region between 40 and 100 feet in medium and low strength materials (Figure 5).

Tunnel closure at the optimum standoff distance (40 to 100 feet) was approximately 50 percent greater in the low strength grout than in the medium strength material; tunnels in the high strength material were virtually undamaged at these standoff distances.

Virtually no closure would be expected in the weaker geologic materials at standoff distances exceeding 250 feet.

A comparison of the model test results with data in real geologic materials (UET granite and sandstone, HARDHAT granite, and NTS tuff) showed that the low strength rock simulant did the best job of modeling tunnel damage (Figure 4); it is believed that the low-strength material best compensates for the joints, cracks, and faults present in natural rock masses.

Limited data from experiments using aluminum tubing to simulate tunnel liners indicates a reduction in damage with increased liner thickness (Figure 5).

A significant increase (approximately 50 percent) in damage to parallel tunnels, as indicated by closure length, occurred when the tunnel centerlines were spaced slightly greater than one tunnel diameter apart (Figure 6).

Closure occurs at a peak free-field radial velocity in excess of approximately 40 ft/sec (Figure 9).

To achieve maximum damage to tunnels 50 feet in diameter and smaller from a 1 KT weapon, the device should be detonated at or above the spring line at a DOB of 40 feet or greater and at a stand-off distance between 40 and 100 feet.

Damage decreases with decreased DOB. Tunnel damage is 80 percent of maximum at DOB of 15 feet. No significant tunnel damage occurs from weapons detonated 15 feet above the surface.

When stemmed or unstemmed devices are detonated inside the tunnel the yield required for closure is a function of tunnel diameter. A weapon yield (KT) $(D/20)^3$ is required for closure with weapons placed inside the tunnel where D is the tunnel diameter.

REFERENCES

1. Department of the Army; "Employment of Atomic Demolition Munitions (ADM)"; Field Manual, FM 5-26, December 1965.

2. Distefano, C. F., and others; "Operation Flint Lock, Shot PILE DRIVER, Project Officers Report - Project 3.1, Response of Tunnel Liners" (U); WT 7010, November 1970, Omaha District, CE, Omaha, Nebraska; SECRET.

3. Holmes, R. S., and others; "Operation Nougat, Shot HARDHAT, Project 3.1, Loading Response and Evaluation of Tunnels and Tunnel Liners in Granite" (U); WT 1801, July 1963, Holmes and Narver, Inc., Los Angell, California; SECRET.

4. Johnson, G. W., Higgins, G. H., and Violet, C. E.; "Underground Nuclear Detonations"; Journal of Geophysical Research, Volume 64, Number 10, October 1959.

5. Engineering Research Associates; "Underground Explosion Test Program, Final Report, Volume II - Rock"; April 1953, St. Paul, Minnesota.

6. Mason, Richard M., and Crossely, Charles H.; "Hasty and Deliberate Tunnel Demolition Tests Conducted Near Scenic, Washington, and Maupin and Madras, Oregon" (U); Report 1408, June 1955, U. S. Army Engineer Research and Development Laboratories, Fort Belvoir, Virginia; FOR OFFICIAL USE ONLY.

7. Hendron, A. J., Jr., Clark, G. B., and Strange, J. N.; "Damage to Model Tunnels Resulting from an Explosively Produced Impulse"; Report 1-6, May 1965, U. S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

8. Joachim, C. E., "Tunnel Destruction, A State-of-the-Art Summary," Masters Thesis Mississippi State University, Mississippi State, Mississippi, December 1977.

9. MacDonald, M. J.; "Superhard Structures in Rock, Nuclear Test Results" (U); DASA 2033, February 1970, Defense Atomic Support Agency, Washington, DC; SECRET.

A COMPARISON OF THE
ACTUAL AND PREDICTED PERFORMANCE OF
A SOLAR ASSISTED SPACE HEATING SYSTEM

DAVID M. JONCICH
CONSTRUCTION ENGINEERING RESEARCH LABORATORY
CHAMPAIGN, IL 61820

I Introduction

Recent increases in fuel prices have resulted in an interest in the development of alternate sources of energy. Solar energy is one such source, as the technical feasibility of heating buildings using flat-plat collectors has been established both in theory and in practice. While the construction phase of solar energy systems requires little more skill than is needed to install conventional heating systems, the design phase is considerably more complex. The fact that solar energy is often not available when it is needed makes system design unconventional in that the sun supplies some, but not all, of the building energy requirement. In addition, an auxiliary source of energy, capable of meeting the peak demand, must be provided to supply heat whenever energy from the sun is not available.

In the past, because of the uncontrolled nature of solar energy, extensive and costly computer studies (taking into account hourly weather data from the site in question) have been required to design and evaluate solar energy systems. For the purposes of determining system feasibility, the cost of these studies can be prohibitive. Thus, there is need for a simple, manual method for making the necessary design calculations and system performance evaluations

prior to a computer analysis. It is the purpose of this paper to describe such a method, and to compare the predictions of the method to the actual performance of a solar facility. Section II of this paper gives a brief description of this manual method, while Section III describes the solar facility from which data was collected. Sections IV and V summarize the system monthly and component performance, respectively. Conclusions follow in Section VI.

II The Universal Curve for Solar Heating

In the course of the development of a solar energy system computer model for use in CERL's Building Load Analysis and System Thermodynamics (BLAST) energy analysis program, CERL performed several hundred hour-by-hour simulations of solar systems used with typical Army buildings in various parts of the country. (1) Analysis of the performance of these systems indicated that, with proper normalization, the performance of a given solar system for all buildings in all locations could be represented for the purposes of feasibility analysis by a single universal performance curve.

A schematic of the type of system under consideration is given in Figure 1. Solar radiation, when available, is converted to thermal energy at the collectors and is transferred to the storage tank by the collector and storage pumps. The heat exchanger, isolating the collection and storage loops, allows the collector fluid to be freezed-protected. Normal operation of the system permits heating of the storage tank whenever sufficient solar energy is available.

Energy is taken from the tank by the load pump (if the storage water is sufficiently hot) whenever there is a demand for heat in the building. If the storage temperature falls too low for heating ($,95^\circ$F), the auto-valve diverts flow around the tank, and the auxiliary heater is energized. As pictured, this system is representative of a large class of liquid solar energy system; all the solar heating simulations run for this study assumed such a configuration.

Use of the universal curve for solar heating, pictured in Figure 2, allows an estimation of the monthly (and seasonal) performance of the system of Figure 1. In order to apply the curve, the user must input only two quantities, the monthly radiation incident on the proposed collector array, and the monthly thermal energy requirement of
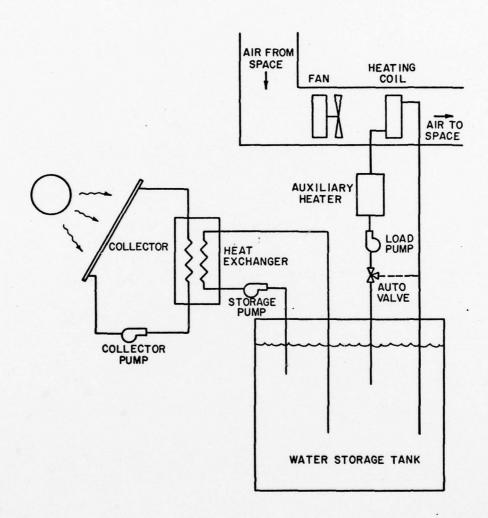
FIGURE 1
CERL SOLAR HOUSE
(SCHEMATIC)

$\rho = 1 - \text{EXP} (-0.46 * P_s ** 0.92)$

MONTHLY INCIDENT RADIATION/MONTHLY ENERGY REQUIREMENTS, P_s

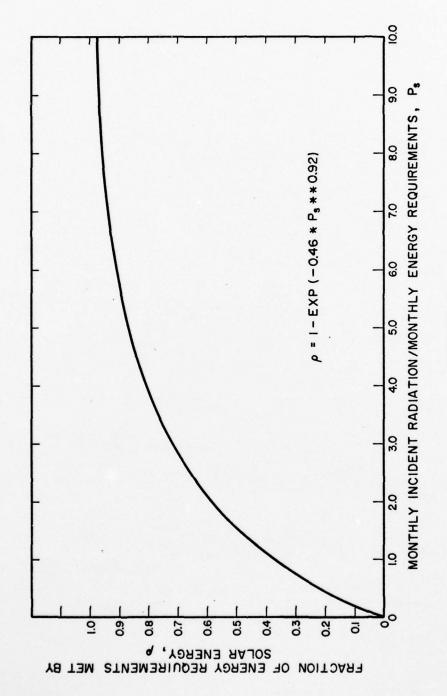FRACTION OF ENERGY REQUIREMENTS MET BY SOLAR ENERGY, ρ

FIGURE 2

THE UNIVERSAL CURVE FOR SOLAR SPACE HEATING

the building in question. The ratio of these two quantities, defined
to be the solar system performance parameter, P_s, is a relative
measure of the amount of available solar energy compared to the ther-
mal energy requirement of the building. Once P_s is computed, the
curve may be consulted directly to estimate the fraction, ρ, of the
building energy requirement which can be provided by the sun. As
derived, it is the function of the universal curve to allow predic-
tion of the monthly solar system performance for any value of P_s. A
more detailed description of the curve (including curves for domestic
heat water and solar heating and cooling applications) is given in
Reference 2.

It can be seen from Figure 2, that the universal curve approach
to solar system design offers a great advantage over an hour by hour
computer analysis. Reasonable values of collector efficiency, heat
exchanger effectiveness, and storage tank heat loss were assumed in
the simulations which produced the universal curve. The fact that
reasonable values for these parameters are contained implicitly
within the curve greatly minimizes the amount of input information
required of the user.

III A Description of the CERL Solar House

Data from CERL's solar house was used to evaluate the accuracy
of the universal curve. This solar house is a 540 sq.ft. residence
which has been retrofit with a solar heating system. Originally
built to test a foam block construction concept, the structure con-
sists of polystyrene blocks 6 inches thick by 12 inches high, and 8
to 10 feet in length. Structural integrity is provided by 3 inch
poured concrete pillars spaced on two foot centers in holes in the
blocks. Because of the thickness of the polystyrene, the thermal
losses of the structure are due almost entirely to infiltration.

The solar system itself (Figure 1) is driven by an array of 12
single glaze, selective surface flat-plate collectors. An inhibited
water-glycol solution is circulated through the array such that each
collector is subject to approximately .63gpm when the collector pump
is active. The collector loop, which contains approximately 12 gal-
lons, is isolated from a pure water storage system by a single pass,
counter flow heat exchanger.

The storage tank is a 1584 gallon precast concrete septic tank which has been foamed with 8 inches of polyurethene insulation and partially buried on the north side of the house. Two self priming pumps draw water from the tank; one (the storage pump) delivering water to the heat exchanger, the other (the load pump) to the heating coil. If, during the heating season, the storage temperature drops below 95°F, a 12KW electric "in line" heater supplies the auxiliary. energy. The auto diverting valve allows heat to be delivered to the coil, while by-passing the tank, when the backup is required.

The control of energy flows within the house is entirely automatic. Collection of solar energy is initiated whenever the collector plate is 10°F warmer than the tank, and is terminated when the collector is within 3°F of the tank temperature. Distribution of the heat to the building is controlled by a room thermostat, in conjunction with a commercially available aquastat (used to determine whether or not the tank is above the 95°F cut-off temperature). When solar energy is available, a demand for heat by the thermostat activates the load pump and distribution fan. As the tank falls below 95°F, the position of the auto-valve is changed (so that the resistance heat is delivered to the heating coil), and the auxiliary heater energized. In this case, the load pump and fan are still in operation.

The house is fully instrumented. A 40 channel data acquisition system records hourly values of solar radiation, fluid flow rates, liquid and air temperatures and energy flows within the system. The performance of the facility for a "typical" sunny winter (20°F) day is given in Figure 3. The instantaneous solar radiation, indicated by the circles on the figure, serves to add energy·to the tank hourly in the amount shown by the squares. This energy increases the tank temperature (with no energy withdrawn) as given by the trianglular plot. Finally, the electrical energy expended by the pumps in collecting the solar energy is also shown, and is seen to be roughly 10% of the energy collected.

IV The Solar System Seasonal Performance

The monthly performance of the CERL solar house was tabulated for February, March, and April of 1977. A summary of the character of the 1977 heating season for these months is given in Table I.
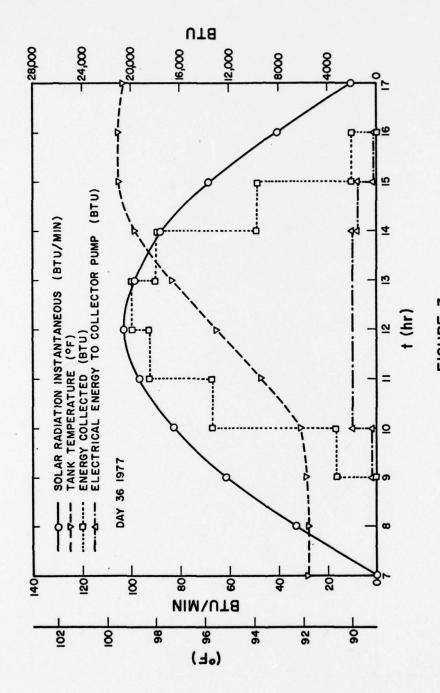
FIGURE 3

TYPICAL SUNNY WINTER DAY SOLAR SYSTEM PERFORMANCE

From the table it is seen that, for the months under consideration, the temperature was generally higher than normal, while the solar radiation was slightly lower.

TABLE I

Temperature (^{O}F)	FEB	MAR	APR
Monthly Mean (1977)	28.6	45.1	57.7
Departure from normal (3)	-.5	+5.7	+6.4

Horizontal Solar Radiation(Btu/ft^2)	FEB	MAR	APR
Monthly Mean (1977)	773	929	1299
Departure from normal (4)	-97	-251	-231

Both quantities required for a computation of P_s were measured directly. The solar radiation was determined by a pyranometer oriented in the plane of the collectors; instantaneous values of this quantity were integrated each hour and summed for the month. The total building load was measured by integrating an $\dot{m}Cp\Delta T$ product of the water delivered to the heating coil (where $\dot{m}$ is the mass flow rate, Cp the fluid specific heat, and ΔT the temperature differential across the coil) and summing this product for the month. The ratio of the solar radiation incident on the collector array to the building thermal energy requirement for each month gave a monthly value for P_s.

Measurement of the fraction, ρ, of the building energy supplied by the sun, was enabled by integrating an $\dot{m}\,Cp\Delta T$ product only when the energy delivered to the heating coil originated from the tank. The ratio of this product to the total load gives the actual percent solar directly.

A comparison of the measured and predicted ρ is given in Table II for each month.

TABLE II

Month (1977)	Slope Radiation (BTUX10^6)	Load (BTUX10^6)	P_s	ρ actual	ρ pred
FEB	6.3	4.2	1.5	.36	.41
MAR	6.9	3.0	2.3	.44	.58
APR	5.1	1.6	3.3	.71	.76

Here, ρ predicted is calculated from the universal curve using the measured value of P_s. For example, in February it is seen that 6.3X10^6 Btu were incident upon the collector array. The thermal energy requirement for this time period was measured at the heating coil to be 4.2X10^6 BTU. The ratio of these quantities gives a P_s of 1.5, which from the universal curve implies a ρ of .41. The actual measurement, however, indicated that the system operated at 36% solar (or ρ = .36).

The general trend in the data is quite evident; the system performance is always lower that expected. Possible reasons for this discrepancy are discussed in Section V.

V Solar System Component Performance

The deviation in ρ measured from ρ predicted led to a more detailed analysis of the solar system component performance, where the three principle components of the solar system are the heat exchanger, the storage tank and the collectors. Data from the performance of these components was compared to the models used in the generation of the universal curve.
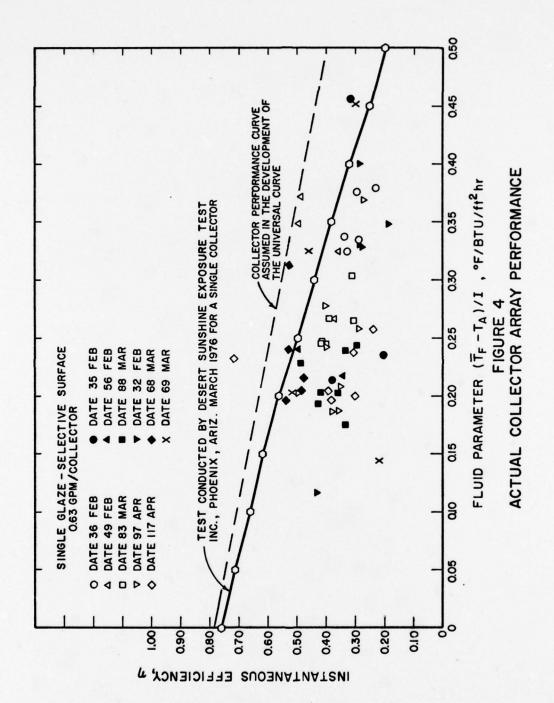
In the derivation of the universal curve for solar heating, a heat exchanger effectiveness of .8 was assumed. An experimental determination of this quantity, is difficult to make because of the small temperature differentials which are present. From measurements of fluid flow rates, specific heats, and temperatures at the inlet and outlets of the primary and secondary sides of the heat exchanger it was found that the heat exchanger effectiveness for the test facility was greater than .8. Since it is shown in Reference 2 that, in this range, the heat exchanger effectiveness has little effect on solar system performance, it was concluded that the heat exchanger

was not causing the discrepancy between the actual and predicted values of ρ.

Heat losses from the storage tank were also examined. Measurements of the tank temperature as a function of time during periods for which energy was neither added to nor removed from the tank indicated that the rate of heat loss from the storage tank was in good agreement with the predictions of the model. (At CERL the decay of tank temperature with time was on the order of $.1^{\circ}F/hr$ under conditions of no energy flow for tanks at $100^{\circ}F$.) Furthermore, independent simulations have shown, that the monthly solar fraction, once again, is not strongly dependent on the rate of heat loss from the tank. According to the model, for a P_s of 2, tank U-values of .05 and .2 $BTU/ft^2-hr-^{\circ}F$ lead to solar fractions of .58 and .56 respectively. This result implies that the tank loss parameter is not responsible for the discrepancy between the measured and predicted monthly solar percent.

The final component to evaluate is the solar collectors. The National Bureau of Standards has defined a procedure for reporting the thermal performance of a solar collector. (5) The results of the test, plotted (open circles) in the N.B.S. format, are given in Figure 4 for a single collector of the type in use at CERL. Here, η is the instantaneous collecter efficiency, (defined to be the ratio of thermal energy collected to incident solar energy), $\overline{T}_F$ the average collector plate temperature, T_A the ambient temperature, and I the instantaneous solar radiation flux. From the figure it is seen that, in spite of the complexity of the collection process, the performance of a solar collector can be described to good approximation by only two parameters, namely the slope and intercept of an N.B.S. plot.

The NBS model for solar collector performance used in the development of the universal curve is shown (dotted line) in Figure 4. This line does not coincide with the actual single collector curve for the CERL collector NBS because the solar system simulations were performed before the performance data for the CERL collectors was available. The discreet data points on the figure show the CERL measurements of the actual performance of the CERL collector array for the dates shown. The general trend in the data is clear; the actual array performance falls short both of the single collector test results and of the line used to generate the universal curve. It is the discrepancy which accounts for a majority of the deviation in ρactual from ρpredicted.

FIGURE 4

ACTUAL COLLECTOR ARRAY PERFORMANCE

While the reason for this degredation of collector array performance is currently under investigation, some comments can be made at this time. The actual collector array is subject to a range of climatic conditions, some of which were not present for the single collector measurement. In particular, collection efficiency is reduced during periods of high winds. Unfortunately, effects of this nature are difficult to estimate quantitatively as a model is required for the "local" wind in the vicinity of the collector.

Other factors exist which complicate the comparison of single collector to array performance. Heat loss in the headers, for example, minimal for a single collector, can be more significant in an array. Furthermore, any effects of long term degradation in plate absorptivity-emissivity or insulation thermal conductivity are not observed during a short duration N.B.S. test.

VI Conclusions

An analysis of the performance of a residential solar heating system was carried out for three months from the 1977 heating season and compared to the performance predicted using the universal curve. The measured solar fraction of energy supplied to the building was found to be less than predicted, but the reduced solar performance was explained, in part, by the fact that the CERL collector array efficiency was consistently lower than the published test results for a single collector. Therefore, in spite of the discrepancy between the actual and predicted percent solar, it is felt that the procedure for solar design offered by the universal curve provides the user with a simple, manual method for making an estimate of the feasibility of a solar heating system.

JONCICH

References

1). D.C. Hittle, D.F. Holshouser, and G.N. Walton, Interim Feasibility Assessment Guidance for Solar Heating and Cooling of Army Buildings, Technical Report E-91 (U.S. Army Construction Engineering Research Laboratory, 1976).

2). D.C. Hittle, G.N. Walton, D.F. Holshouser, D.J. Leverenz, Predicting the Performance of Solar Energy Systems, Interim Report E-98 (U.S. Army Construction Engineering Research Laboratory, 1977).

3). from Illinois State Water Survey

4). from Climatic Atlas of the United States (National Oceanic and Atmospheric Administration, 1968).

5). J.S. Hill and T. Kusauda, Methods of Testing for Rating Solar Collectors Based on Thermal Performance, NBSIR-74-635 (National Bureau of Standards, 1974).
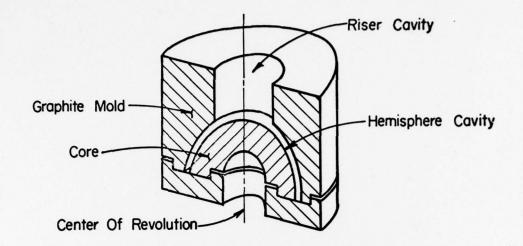
USE OF COMPUTERS IN MOLD DESIGN (U)

*ROBERT E. KEENAN, JR., MAJ, EN
MR. WILLIAM C. ERICKSON
LOS ALAMOS SCIENTIFIC LABORATORY
LOS ALAMOS, NEW MEXICO 87545

Mold design is an art. Numerous design techniques are used
to ensure that the casting obtained from a mold is sound. Normally,
the designer uses a trial-and-error method to perfect his mold. If
defects in the casting are discovered, he changes the design and casts
another part. As the size and complexity of the mold increase, how-
ever, this approach becomes economically unfeasible, especially when
the cost of the metal to be cast is high.

At the Los Alamos Scientific Laboratory, it is often nec-
essary to cast parts of uranium, gold or other metals that are rare
and/or of high chemical and isotopic purity. Uranium parts are cast
in zirconium-oxide-coated graphite molds which are relatively expen-
sive to fabricate. Also of significance is the recovery cost; i.e.,
the cost to repurify the extra metal after the casting has been ma-
chined. These costs are important in two ways. First, they provide
motivation to design better molds so that sound castings can be ob-
tained with less waste. Second, they prevent the use of the ordinary
trial-and-error method of mold design.

This paper presents our efforts to determine a system to
evaluate mold designs using computer simulations. Although we report
on the mold design for a thin-walled uranium hemisphere, the system is
applicable to virtually any shape mold.

Cross Sectional View

Figure 1. Current mold design for hemispherical casting.

BACKGROUND

Figure 1 depicts an example of a mold used to cast a thin-walled uranium hemisphere.

An allowance for machining is included in the hemisphere cavity, and another large volume of "waste" metal is in the riser cavity. The riser performs two essential functions in this design. First, it provides a reservoir of molten metal to feed the casting as it shrinks during solidification. Second, it moves the thermal center of the system from the part into the riser. As a result, a temperature gradient is created which allows the casting to cool from the equator towards the riser.* Solidification occurs in the vicinity of

---

*It is important to note that this temperature gradient occurs only if the initial temperature of the mold is below the melting point of the molten metal.

all the interior surfaces, but not rapidly enough to close off any still-molten part from the riser. In a typical mold designed on these principles, the riser and machine allowances contain more metal than is in the final machined hemisphere.

When casting hemispherical shapes in chill molds, another phenomenon is of interest. Figure 2 is a picture (taken with a flash x-ray machine) of a mold as it is being filled with lead. As the molten metal hits the mold core (A) it is splashed against the outer
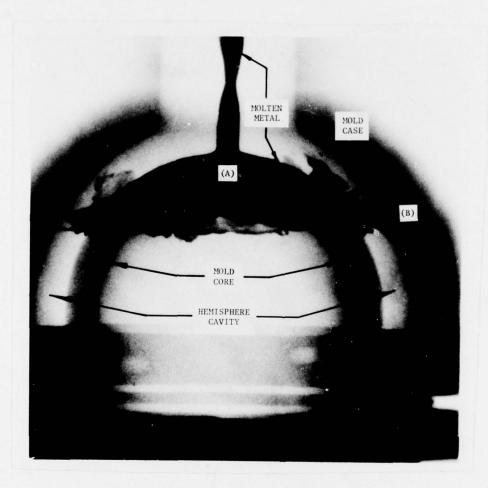

Figure 2. Flash radiograph of molten metal entering mold.

cavity wall (B). Because the mold is below the melting temperature of the metal, some freezing of the metal occurs where it strikes the outer wall. Because of the slashing, voids are frozen into the metal. In thin-walled castings, the metal cools so quickly that this metal never remelts to free the voids as the mold is filled. In addition, this area acts as a trap for gases rising from the lower part of the casting. Consequently, additional allowance must be made to insure that these defects can be machined away.

With the objective of reducing or eliminating all of these problems and thereby reducing costs, we decided to design a new mold. We reasoned that if we could preheat the mold above the melting temperature of the metal, then withdraw heat from the system at the equator of the casting with some type of cooling system, we could obtain a temperature gradient in the casting and mold which would force the solidification front to move in the desired direction. Since the initial temperature of the mold would be above the melting temperature, solidification would not occur along the mold surfaces. This would allow us to reduce the machine allowances. The riser would stay molten longer, allowing us to reduce its size as well.

To evaluate a new design, we need to know how the solidification front advances as the casting cools. For thin-walled shapes, it is impossible to obtain this information experimentally. Molten uranium will react with bare thermocouples. Thus the thermocouples could not be placed in the molten metal and would either have to be inserted into the mold or encased in ceramic. Both of these situations cause inaccuracies for which it is very difficult to account. Therefore, to evaluate our directional solidification concept, we decided to simulate the cooling process using a digital computer.

DISCUSSION

Computer Programs

For heat-transfer calculations, we used the Chrysler Improved Numerical Differencing Analyzer for 3rd Generation Computers (CINDA-3G).[1] It is a computer program in general use at LASL.* This program employs the user's choice of several finite differencing algorithms to obtain solutions to transient or steady state thermodynamic systems.

---

*CINDA-3G was developed by the Chrysler Corporation Space Division under contract NAS9-7043 to the National Aeronautics and Space Administration's Manned Spacecraft Center at Houston.

*KEENAN & ERICKSON

The thermal system is presented to CINDA-3G in a network format, consisting of nodes and conductors. Each node is identified by number, volume, and a reference to an array of its thermal properties. Each conductor (one between each pair of adjacent nodes) is identified by number, two node numbers, and a reference to an array containing temperature-varying thermal conductivity.

Three types of nodes are allowed. Diffusion nodes have the capability to store energy (thermal capacitance). Arithmetic nodes have no thermal capacitance, and negligible volumes. They are used to obtain temperatures at specific points of interest. Boundary nodes are used to set mathematical boundary conditions. Temperatures of boundary nodes are not changed by the differencing subroutines, but may be changed by the user during a simulation.

Only two types of conductors may be used, regular and radiation. Thus, a simple one-dimensional network might look like figure 3. In this system nodes and conductors are of the types indicated in Table I.

Capacitance (C) is defined only for diffusion nodes. It is calculated as the product of node volume and the temperature-dependent properties of density and specific heat.
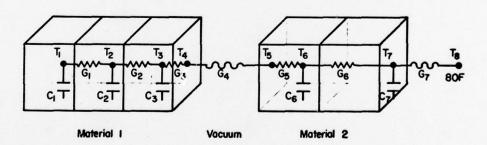
$$C_1 = V_{i\sigma}c_{pi}$$

Figure 3. Example network of 1D thermal system.

*KEENAN & ERICKSON

TABLE I

NODE AND CONDUCTOR TYPES FOR SYSTEM IN FIGURE THREE

Node	Type	Conductor	Type
T1	Diffusion	G1	Regular
T2	Diffusion	G2	Regular
T3	Diffusion	G3	Regular
T4	Arithmetic	G4	Radiation
T5	Arithmetic	G5	Regular
T6	Diffusion	G6	Regular
T7	Diffusion	G7	Radiation
T8	Boundary		

Conductance (G) calculations depend on the type of conductor. Through solid material,

$$G_i = \frac{kA}{x}$$

where   k = temperature-dependent thermal conductivity
        A = cross-sectional heat flow area
        x = path length between node centers.

For convective heat transfer,

$$G_i = hA$$

where   h = convective heat transfer coefficient
        A = cross-sectional heat flow area.

In the case of radiative heat transfer, conductance is calculated using this equation:

$$G_i = \sigma A F$$

where σ = Stefan-Boltzmann constant

        A = radiant surface area

        F = net radiant interchange factor.

The extension of this one-dimensional system into two or three dimensions is not difficult to visualize. However, if done by hand, the calculation of node volumes, distances between node centers and cross-sectional heat flow areas would be tedious and time-consuming. For axiosymmetric molds, these calculations are handled by FED-JB,* a program which also writes a disk file containing the node and conductor data. After program control constants, thermal property arrays, and other instructions are added, this file is used as the input data for CINDA-3G.

For molds containing hundreds of nodes, a large volume of data is generated. Even if only the data for the uranium nodes were printed, temperatures at every time step would fill hundreds of pages of output. In addition, the manipulation of this data by hand to find meaningful results would be extremely time-consuming.

Using a set of subroutines previously developed by the authors at LASL, the data can be presented on 16 mm or 35 mm color film.[3] At the end of each specified time interval, a frame is plotted showing each node of the network. One of several available characters (for instance, *, +, -) is plotted on the film at a point corresponding with the center of each node. The color of each character is set according to the temperature range within which the node falls. When 16 mm film is made, the temperature changes can be studied as the film is projected as a motion picture. This allows rapid determination of the effects of changes in mold design, since the advancement of the solidification front can readily be seen.

Design Procedures and Results

The initial concept of a cooling system of a hemispherical mold is shown in figure 4.

For computer simulation, a very simple model was constructed to first test the feasibility of the directional solidification concept. Assumptions were: (1) that the mold cavity is filled instantaneously; (2) that intimate contact exists between metal and mold. For simplicity, the mold coatings (normally flame-sprayed zirconium

---

*FEDJB is the result of modifications by Jim Burns, Group WX-1, LASL, to FED.[2] FED is a "heat mesher" developed at Lawrence Livermore Laboratory (LLL) to calculate the same type of input for TRUMP, another computer program for thermodynamic problems.
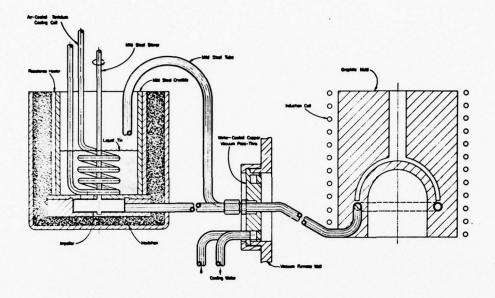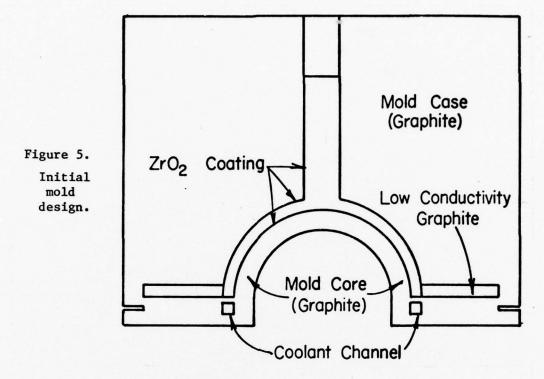
Figure 4. Schematic of cooling system for hemispherical mold.

dioxide) were ignored. The initial temperature of the melt was 1275°C
(2327°F) and that of the mold was 1135°C (2075°F). An arbitrary 500
Btu/min was removed from the system from the node directly under the
equator of the casting.

The results of this initial model proved the feasibility of
the concept. They also indicated that the transfer of heat through
the graphite conductors to the heat extraction area must be reduced.
We then proceeded to design our mold.

Our initial mold design incorporated a square channel ma-
chined into the graphite mold through which liquid tin at 350°C (662°F)
would circulate as a coolant. This channel is separated from the cast-
ing by a thin sheet of graphite (see figure 5). The zirconium dioxide
mold coatings are included and a piece of low-conductivity graphite is
inserted to reduce the heat flow through the mold. For this simula-
tion, the center tin node is kept at 350°C throughout. Otherwise, the
assumptions are the same as for the previous model.

Figure 5.

Initial
mold
design.

Results of this simulation indicate that the core cools off much more rapidly than the case, causing the uranium to solidify far ahead of the outside. This is undesirable because a portion of the hemisphere may be frozen off from the riser, causing voids to be formed. Figure 6 shows how the solidification front advances.

In our next design there are four changes: First, to obtain better heat transfer from the uranium to the coolant, we inserted a tantalum tube to carry the molten tin. The uranium is assumed to be in intimate contact with the tantalum.
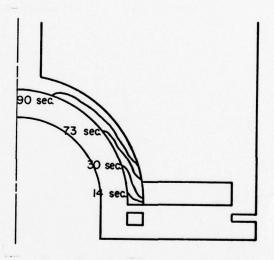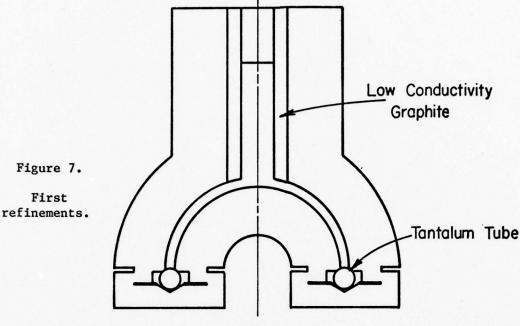

Figure 6. Results of simulation.

Figure 7.

First
refinements.

Low Conductivity
Graphite

Tantalum Tube

Second, the graphite channel in which
the tantalum tube rests is coated
with zirconium-dioxide, a low-conduc-
tivity flame-sprayed material, to re-
duce the heat flow from the graphite
to the coolant. Third, the low-con-
ductivity graphite is moved to the
riser to help maintain its tempera-
ture. Lastly, the mass of the mold
is reduced on the outside and in-
creased in the core. This is to
help balance the temperature gra-
dient between core and case, there-
by bringing the liquid-solid inter-
face in the metal more perpendicu-
lar to the mold walls. Figure 7
shows the resultant design.

Results of the simu-
lation (figure 8) indicate that we
still have not done enough to force
the heat to flow through the metal
to the coolant. The core still

240 sec

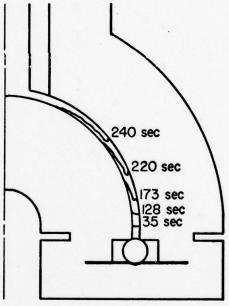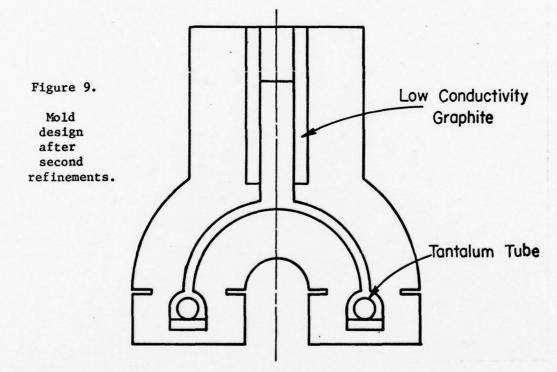220 sec

173 sec
128 sec
35 sec

Figure 8.
After first refinements.

cools off too soon in comparison to the mold case, causing the solidi-
fication front to advance rapidly along the core.

A study of heat flow rates through specific nodes in the
vicinity of the tantalum tube shows that a large amount of heat is
still being transferred from the mold to the coolant tube through the
small contact areas at the bottom and top of the tube. Additionally,
it is likely that in an actual casting, the uranium would shrink away
from the tantalum upon solidification. This would cause a large drop
in thermal conductivity at the interface.

Our next design includes the following modifications (see
figure 9):
1. A layer of a castable ceramic is placed under the tanta-
lum tube. This low-conductivity material is to reduce heat flow
through the bottom of the tube.
2. The tantalum tube is lowered to allow molten uranium to
flow around it. This accomplishes two things: first, the contact be-
tween the top of the tube and the mold is eliminated; second, the ura-
nium, upon solidification, will shrink around the tube, reducing the
interface problem.

Figure 9.

Mold
design
after
second
refinements.

Low Conductivity
Graphite

Tantalum Tube

The computer simulation after these second refinements shows significantly better results. The solidification front advances perpendicular to the mold walls until the hemisphere is almost entirely solidified. When the metal finally does solidify across the core, the rest of the front is not far enough behind it to cause any feeding problems. See figure 10.

Using this final design we estimate that we can obtain, assuming 0.03-in. machine stock, a hemisphere of 2.30-in. inside radius, and 2.40-in. outside radius. The riser has a radius of 0.5 in. and a height of 1.5 in. Table II compares the new mold design with one that would be required to cast the same part using current methods.

The right column in Table II includes figures for a new design which, as this is written, is being prepared for simulation. The new design (figure 11) is an attempt to reduce the amount of uranium around the cooling tube. The tantalum tube is moved up to alongside the base of the hemisphere. Since the coefficient of expansion for tantalum is less than that for uranium, the tube must be placed in the mold core.
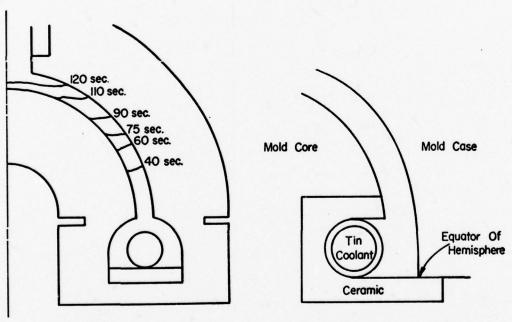
Figure 10.
After second refinements.

Figure 11.
Future refinements.

*KEENAN & ERICKSON

TABLE II

DIMENSIONS OF NEW AND OLD MOLD DESIGNS

	Finished Part Dimens. (in.)	(in.³)	Current Design Mold Dimens. (in.)	(in.³)	Directional Solidification Mold Dimens. (in.)	(in.³)	Proposed Simulation Dimens. (in.)	(in.³)
Hemisphere								
inside radius	2.30		2.22		2.27		2.27	
outside radius	2.45		2.57		2.48		2.48	
volume		5.32		12.64		7.44		7.44
Riser								
radius			1.0		0.5		0.5	
height			2.0		1.5		1.5	
volume				6.28		1.18		1.18
Extra Volume								
at equator		___		1.32		8.96		0.61
Total Volume		5.32		20.24		17.58		9.23
Volume Reduction						2.66		11.01

Otherwise, the uranium would shrink away from the tube. Since this works against the requirement to reduce the heat transfer through the core, the tube is encased in a low-conductivity ceramic open only to the uranium. This should direct the majority of the heat flow through the uranium. Also, the mass of the core is increased. The results of this new simulation are not available at the time of this writing. However, previous simulations indicate that we can expect this new design to give better solidification front propagation and a reduced volume of uranium in the coolant area.

CONCLUSIONS

Directional solidification of the hemispherical casting has been attained. A further reduction of the heat flow through the core would be beneficial and should be attained in the next design. This further reduction, however, does not appear essential to sound castings.

*KEENAN & ERICKSON

The final simulated mold design saves 2.66 cubic inches of metal. For uranium, that is 817 grams (28.8 ounces) saved. If the new design works as expected, the savings will be 11.01 cubic inches, which is almost 3.4 kilograms (7.45 pounds) These savings are significant.

Finally, we have developed a system for improving the design of molds where cost and/or complexity makes the trial-and-error method unfeasible. This system is by no means limited to thin-walled hemi-spherical castings. Since CINDA-3G has a three-dimensional capability, virtually any shape mold can be simulated and redesigned.

References

1. Lewis, D. R., Gaski, J. D., and Thompson, L. R., "Chrysler Improved Numerical Differencing Analyzer for 3rd Generation Computers," Technical Note TN-AP-67-287, 20 October 1967.

2. Schauer, Dale A., "FED: A Computer Program to Generate Geometric Input for the Heat-Transfer Code TRUMP," UCRL-50816, Rev. 1, 12 January 1973.

3. Erickson, William C., "The Use of General Purpose Software in Casting Solidification Simulation," Metallurgical Transactions B, Vol. 813, March 1977, pp 93-97.

THE MECHANISM OF CORROSION INHIBITION
BY *DINONYLNAPHTHALENESULFONATES*

PAUL J. KENNEDY
US ARMY MOBILITY EQUIPMENT RESEARCH AND DEVELOPMENT COMMAND
ENERGY AND WATER RESOURCES LABORATORY
FORT BELVOIR, VIRGINIA 22060

I   INTRODUCTION

Operability of combat and combat support equipment can be seriously compromised by rusting of power train components. As a preventative measure, most lubricants in the Army Supply System are rust inhibited. Sulfonates, such as the petroleum sulfonates and salts of *dinonylnaphthalenesulfonic acid*, have proven to be effective rust preventatives for Army lubricants, however, their use is limited due to their high temperature instability and viscosity. New rust inhibitor systems are needed for formulation into a variety of fluids such as turbine engine oils, aircraft hydraulic fluids and the newer synthetic nonhydrocarbon fluids. A basic understanding of the rust inhibition mechanism is needed in order to accomplish this task and to keep pace with potential rust problems associated with advanced Army systems.

Most, although not all investigators, believe that rust inhibition is the result of rust inhibitor adsorption on metal surfaces. Adsorption of a monomolecular layer on the metal surface could act as a barrier toward diffusion of corrodants such as water and oxygen to the surface and oxidation products from the surface (1). Our previous work (2) on the mechanism of rust inhibition indicated that sulfonates are strongly and irreversibly adsorbed on metal oxide surfaces and form close-packed monolayers at low solution concentrations. These findings support a barrier-type mechanism for corrosion inhibition, however, the actual mechanism is probably of much greater complexity. The effect of micelles in bulk solution, co-adsorption of the base lubricant, competitive effects of other additives in the lubricant system and chemical modification of the base metal by adsorption needs to be considered.

Our present studies are directed towards a better understanding of the rust inhibition mechanism through an investigation of the thermodynamics associated with the chemical processes that occur in a typical rust inhibiting system. The first part of this investigation was concerned with the competitive adsorption of water with sulfonates on iron oxide in order to better understand the corrosion process. The second part of this investigation is concerned with measuring the heat liberated by sulfonate adsorption since this is a measure of the extent of interaction between the sulfonates and the metal oxide substrate.

II   EXPERIMENTAL SECTION

Materials used in this investigation included sodium dinonyl-naphthalenesulfonate (Na(DNNS)) and barium dinonylnaphthalenesulfonate (Ba(DNNS)$_2$); cyclohexane (J. T. Baker Chemical Company, Instra-Analyzed Grade) was used as received and also as a "dry" solvent by treatment over 4Å molecular sieve. The iron (III) oxide powder was of relatively high purity and had a surface area of 28 m^2/g as determined by the BET gas adsorption method. The method of preparation for the sulfonates was previously described (2).

Information on the mechanism of corrosion inhibition was obtained by use of a flow microcalorimeter (Microseal Ltd., England). The principal of operation of this flow calorimeter is very similar to that of a column chromatograph except that thermistors are inserted in the adsorbent bed to detect temperature changes due to either adsorption or desorption of material from the column. In our experimental work, the iron oxide powder is used to form the adsorbent bed. In the continuous flow mode of operation, the carrier liquid (pure cyclohexane) is pumped through this miniaturized bed by special mircrpumps at a flow rate of 0.05 ml/min. so as to establish temperature equilibrium. After establishment of temperature equilibrium, a dilute solution of sulfonate in cyclohexane is percolated through the adsorbent bed until the heat effect associated with adsorption of sulfonate on iron oxide powder is dissipated. The flow of the pure cyclohexane carrier liquid is then resumed in order to measure any heat effects associated with the desorption of sulfonates. In the pulsed adsorption mode of operation, small amounts of dilute sulfonate solution are periodically injected directly into the carrier liquid stream to determine associated heat effects.

The heat adsorption of Na(DNNS) on iron oxide was determined by using a thermometric titration and solution calorimeter (Tronac, Inc.). Heat of adsorption measurements were made by titrating a suspension of iron oxide powder in cyclohexane with a dilute solution

of Na(DNNS) in cyclohexane at 30°C using either a 25 ml or 50 ml isothermal reaction vessel. This data was augmented by ultraviolet analysis of the post run supernatant liquid to determine the extent of sulfonate adsorption on the iron oxide powder.

III RESULTS AND DISCUSSION

Water is a common contaminant in lubricating oils and would be expected to actively compete for adsorption with the other lubricant additives. The effect of water can be complex; it can either increase adsorption, decrease adsorption or have little or no effect on the adsorption of other additives. For example, we had found that water adsorbed on nickel oxide can act as sites for additional adsorption of phenol (3) or it can effectively block the adsorption of other additives such as phenyl disulfide. In the case of phenyl disulfide, room temperature adsorption studies indicated an adsorption of only 2 to 4 µg of disulfide per gram of iron oxide as received. Moderate drying of the iron oxide powder at 120°C under vacuum conditions resulted in an increased adsorption of 24 µg/g. Intensive drying at 350°C resulted in a further increased adsorption to 172 µg/g. In this latter case dehydroxilation of iron oxide surfaces with generation of highly active sites not present on the moderately dried sample were suspect. In the previous work (3), we had found that sulfonate adsorption isotherms were not noticeably effected by water and first studies were initiated without consideration for the effect that drying or the iron oxide powder would have on experimental results.

Effect of Water

Data output from the flow calorimeter is similar to the output from a gas chromatograph in that detector response is recorded as a function of time. A typical recorder output from the flow calorimeter is shown in Figure I. Initially, a steady flow of cyclohexane is percolated through the iron oxide column in order to achieve a steady thermal base line and at this point, a flow of dilute Ba(DNNS), in cyclohexane is substituted for the flow of cyclohexane. The thermal effect produced by the adsorption of BA(DNNS)$_2$ on the column is recorded as a pen deflection corresponding to region "A" in Figure I. (Exothermic processes produce a pen deflection in the -Q direction and endothermic processes produce a pen deflection in the +Q direction.) The area under the curve is directly proportional to the heat of adsorption. Flow of the sulfonate solution is continued until the recorder returns to the initial thermal base line. The flow of sulfonate is then discontinued and replaced by a flow of pure cyclo-

hexane which allows the measurement of thermal effects associated with desorption. This cyclohexane wash corresponds to region "B". Regions "C" and "D" correspond to additional flow change-overs to $Ba(DNNS)_2$ in cyclohexane and pure cyclohexane, respectively.

Data initially obtained from the flow calorimeter (Figure I) using $BA(DNNS)_2$ as an adsorbate and iron oxide powder for the adsorbent bed was at first somewhat puzzling. Adsorption of sulfonate would be expected to produce an exothermic heat effect and further washing with pure cyclohexane should produce a very slight endothermic heat effect due to slight desorption of sulfonate. These types of heat effects were initially not found, in fact, the reverse of these heat effects were observed. The initial work indicated that sulfonate adsorption produced a slight exothermic heat effect as would be expected for sulfonate adsorption, however, this was rapidly over-shadowed by a very much greater endothermic heat effect of unknown origin. Washing with cyclohexane produced an exothermic heat effect suggesting an adsorption process rather than a desorption process. Further addition or sulfonate produced an endothermic heat effect of approximately the same magnitude as was found during the previous cyclohexane wash.

The cause of these unusual thermodynamic effects was soon discovered. The solubility of water in cyclohexane at room temperature is approximately 15 ppm and the reagent grade cyclohexane used in this work contained 10 ppm of water. Thus, the initial flow of "wet" cyclohexane over the iron oxide powder column that was necessary to establish temperature equilibrium also produced a water saturated iron oxide surface. This suggested a possible explanation for the unusual heat effects shown in Figure I. Region "A" can be interpreted as indicating an exothermic sulfonate adsorption process that is over-shadowed by an endothermic heat effect associated with desorption and solubilization of water by sulfonate micelles. Washing the sulfonate monolayer with "wet" cyclohexane as in region "B" indicates an exothermic effect that can only be associated with water adsorption through the sulfonate monolayer. (It had previously been shown that concentrations of $Ba(DNNS)_2$, much less than 0.5%, produce close-packed monolayers.) Further addition of sulfonate as in region "C" produces an endothermic heat effect that again can be attributed to solubilization of adsorbed water by the micelles.

This possible explanation was then investigated by studying the effect that drying of the cyclohexane and sulfonate solutions had on the heat of sulfonate adsorption. Initial work involved the use of cyclohexane dried over 4Å molecular sieve for establishment

of thermal equilibrium and subsequent washing. No attempt was made to dry the sulfonate solution for this part of the study. This data is shown in Figure II. Region "A" indicates an exothermic adsorption, region "B" indicates no desorption or an extremely slow desorption, but region "C" which corresponds to a re-wash of the adsorbed sulfonate monolayer with sulfonate indicates an exothermic heat effect. This was again attributed to water adsorption. This was verified by use of both dried solvent and dried sulfonate solution. These last studies indicated an exothermic heat effect in region "A" and no heat effects in region "B", "C", and "D". This particular recorder tracing is not shown since it is identical in shape to Figure II except that no heat effect was observed in region "C".

Table I quantitates the effect of water on sulfonate adsorption and represents data obtained under the above described conditions. It is interesting to note that the net heat effect obtained by summing the heats of adsorption and desorption is a constant independent of run conditions.

TABLE I. Effects of Water on the Adsorption of Sulfonates

Run No.	Heat of Adsorption for Na(DNNS) on Fe_2O_3 (mcal)	Heat of Desorption from Fe_2O_3 (mcal)	Net Heat Effect (mcal)
A. "Wet" Cyclohexane/"Wet" Sulfonate Adsorption			
1	-16.9, +2.4	-8.6	-23.1
2	-21.7, +5.7	-8.2	-24.2
3	-22.0, +5.1	-5.8	-22.7
4	-19.6, +1.8	-5.1	-22.9
B. "Dry" Cyclohexane/"Wet" Sulfonate Adsorption			
5	-58.2	+---	---
6	-62.0	+42.7	-19.3
7	-55.0	+---	---
C. "Dry" Cyclohexane/"Dry" Sulfonate Adsorption			
8	-22.7	N.D.	-22.7
9	-23.3	N.D.	-23.3
10	-23.2	N.D.	-23.2
11	-22.5	N.D.	-22.5

These unusual results indicate some of the complexities involved in the rust inhibition mechanism. It had previously been reported that sulfonates are strongly and irreversibly adsorbed to form close-packed monolayers at solution concentrations significantly less than used in this work (2). The irreversible nature of sulfonate adsorption is verified by results under condition "C" of Table I. This strongly suggests that water is the chemical species that is being either adsorbed or desorbed under the other test conditions. For Run Nos. 1-4, we are dealing with an initially water saturated surface and the endothermic heat effect is most probably the result of water solubilization in the sulfonate micelles. In the second series of experiments (Run Nos. 5-7) we are dealing with an initially dry surface and the large exothermic heat effect suggests co-adsorption of water with the sulfonate. This type of behavior could be explained by assuming that an equilibrium exists between adsorbed water and micellar water and that the direction of this equilibrium is determined by the amount of adsorbed water initially on the metal surface. Thus, in a dry system, water solubilization in the sulfonate micelle might be considered as the first step in the corrosion inhibition mechanism.

Another factor that needs to be considered is the unusual shape of the sulfonate molecule. This molecule is umbrella shaped and is quite bulky (ca. 100 A^2/molecule) and would thus be expected to interact with only a relatively few of the metal oxide sites capable of adsorption even in keeping with the formation of close-packed monolayers. This would explain the ability of small water molecules to penetrate the sulfonate monolayer and adsorb on the metal oxide surface. It is reasonable to assume that these sites would be available to interact with the hydrocarbons of the bulk lubricant and would thus explain the influence of the structure of the bulk lubricant on the mechanism of corrosion inhibition.

The effect of trace amounts of water on the adsorption process is illustrated in Figure III. In this work, Na(DNNS) in cyclohexane is stored over molecular sieve in order to prevent water adsorption from the atmosphere. The syringe used to inject the Na(DNNS) into the carrier stream is usually freshly loaded before the start of the run. In Figure III, the first four injections were from a Na(DNNS) solution that was allowed to stand overnight in the syringe. After the fourth injection of sulfonate, fresh Na(DNNS) solution was drawn into the syringe and used for the remaining injections. It appears that the amount of water in the micelle, the degree of dryness of the cyclohexane and the amount of water on the metal oxide surface determine to a great extent the "apparent" heat effect associated with adsorption. This complex water equilibrium is probably the key to the understanding

of why sulfonates are effective in some base fluids and not in others.

Heat of Adsorption

Analysis of data obtained by thermometric titration of suspended Fe_2O_3 in cyclohexane with Na(DNNS) can be difficult since total heats evolved (~0.3 cal) are small compared to other potential side reactions. A program has been written to reduce the effect of any initial spurious heat effects and to facilitate data reduction. Typical titration data as obtained from the solution calorimeter is shown in the bar graph (Figure I). The curved line in this graph represents the heat that would be evolved for a theoretical Langmuir adsorption with ΔH, K_2 and "a" equal to 9 Kcal/mole, 1.05×10^4 liter/ mole and 0.32 μ mole/$\bar{m}^2$, respectively. The fit of the theoretical curve to all parts of the bar graph is good. Although the adsorption constants obtained from the solution calorimeter indicate strong adsorption and close-packing, these values are somewhat lower than previously reported (2).

Initial work indicated a relatively large uncertainty in the measured value for the heat of adsorption of sulfonates on iron oxide powder. In order to reduce this uncertainty, work was done on the effect of various parameters (e.g., the weight of iron oxide powder and effect of water) on the measured heat of adsorption. Heats of adsorption measurements on iron oxide powder dried under vacuum at $120^{\circ}C$ were not highly reproducible. On the other hand, experiments with iron oxide powder that were not subjected to drying gave very reproducible results. However, the measured heat of adsorption decreased with the weight of iron oxide powder. This effect can be seen in Figure V.

The graph shows the amount of heat liberated (mcal) during the titration of various weights of iron oxide powder in 25 ml of dry cyclohexane with a 0.0321 m solution of sodium dinonylnaphthalene-sulfonate. In Figure V, the solid lines represent actual data points and the dashed lines (e.g., 1.0 and 0.5 g run) are straight line extensions of the initial slopes. At the start of these thermometric titrations, nearly all sulfonate molecules are adsorbed. As the surface coverage increases, the fraction of sulfonate molecules that adsorb on the iron oxide powder decreases and this results in a general falling off of the heat evolved with respect to ml titrated. This can be seen in Figure V for the 1.0 and 0.5 g samples. If all molecules were adsorbed, then the experimental points would be expected to lie on the dashed lines. The initial slope in this type of plot (heat evolved/amount adsorbed) would be equal to the heat of adsorption and should be completely independent of the amount of iron oxide

7

KENNEDY

powder used in the experiment. This is obviously not true in this experimental work. Calculated heats of adsorption were 8.5, 8.2, 7.8, 7.1, 3.9, and 2.1 Kcal/mole for 2.5, 2, 1.5, 1, 0.5, and 0.25 gram samples.

This unusual effect can be attributed to an endothermic heat effect (associated with water desorption) that occurs along with the exothermic heat effect associated with sulfonate adsorption. According to competitive adsorption theory, the concentration of water and sulfonate in solution are dependent on the fraction of the total surface covered by sulfonate molecules and this would fully account for this unusual effect. This hypothesis can be tested by assuming that the actual measured amount of heat (Q_m) is equal to the exothermic heat of adsorption (Q_a) and the endothermic heat associated with water desorption (Q_d).

$$Q_m = Q_a - Q_d \tag{1}$$

Dividing both sides of equation (1) by grams of adsorbent (g), we have:

$$Q_m/g = Q_a/g - Q_d/g \tag{2}$$

If equation (2) is valid, a plot of Q_m/g vs $1/g$ should give a straight line having an intercept of Q_a/g and a slope of Q_d. A plot of experimental data according to equation (2) is shown in Figure VI for a surface coverage value of 19.3 µ moles/g. This plot indicates that the true heat of sulfonate adsorption is 10.1 Kcal/mole.
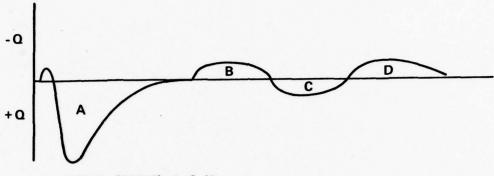
IV   CONCLUSIONS

Previous studies (2,3) have indicated that sulfonates are strongly adsorbed on metal oxide powders. This work supports this previous finding in that a large heat of adsorption was found (10.1 Kcal/mole) for sulfonate adsorption on iron oxide powder. This indicates a relatively strong interaction between sulfonate and the metal oxide surface. This work also pointed out other areas that need consideration for possible interpretation of the mechanism of corrosion inhibition. Other types of rust inhibitors (e.g., long chained organic acids) have molecular dimensions that approximate the spacing of sites on the metal oxide surface, however, salts of dinonylnaphthalenesulfonic acid have an unusual umbrella-like shape and are quite bulky by comparison having molecular areas ranging from $200A^2$ to $125 A^2$ at 40 dyn/cm (4,5). Thus, sulfonate molecules will interact with only a relatively few of the sites available for

adsorption. This would explain the ability of the small water molecules to penetrate the sulfonate monolayer and adsorb on these free sites. It appears that in dry systems, water absorption in the micelle is favored over adsorption on the metal surface and that this factor would also tend to inhibit the corrosion process. The cyclohexane molecule is apparently too sterically hindered or lacking in adequate polarity to interact with these sites. A facile equilibrium between surface water and micellar water was indicated. In actual corrosion inhibiting systems, the solvent is normally an oil containing polar components and it would be expected that these polar components would be capable of interacting with these "free" sites which would tend to block this water equilibrium process.

REFERENCES

1. R. R. Annand, R. M. Hurd, and N. Hackerman, J. Electrochem. Soc., 112, 138 (1965).

2. P. Kennedy, M. Petronio, and H. Gisser, J. Phys. Chem., 74, 102 (1970).

3. P. Kennedy, M. Petronio, and H. Gisser, Ibid., 75, 1975 (1971).

4. M. L. Smith, B. E. Gordon, and R. C. Nelson, Ibid., 69, 3833 (1965).

5. F. M. Fowkes, Ibid., 66, 1843 (1962).

A   0.5% Ba (DNNS)$_2$ in C$_6$H$_2$
B   C$_6$H$_{12}$ with 0.01% H$_2$O
C   0.5% Ba (DNNS)$_2$ in C$_6$H$_{12}$
D   C$_6$H$_{12}$ with 0.01% H$_2$O

Figure I. Adsorption of Ba (DNNS)$_2$ on Fe$_2$O$_3$ with "wet" C$_6$H$_{12}$

A   0.5% Ba (DNNS)$_2$ in C$_6$H$_{12}$
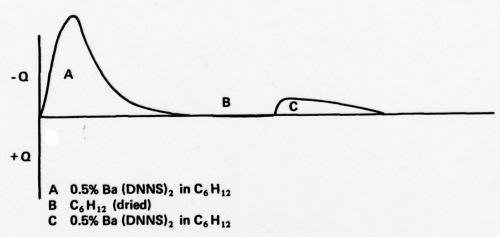B   C$_6$H$_{12}$ (dried)
C   0.5% Ba (DNNS)$_2$ in C$_6$H$_{12}$

Figure II. Adsorption of Ba (DNNS)$_2$ on Fe$_2$O$_3$ with "dry" C$_6$H$_{12}$
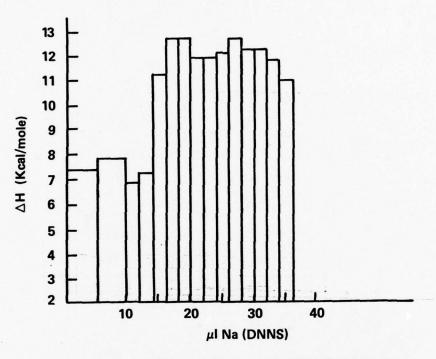
Figure III. Effect of Water on Na (DNNS) Adsorption
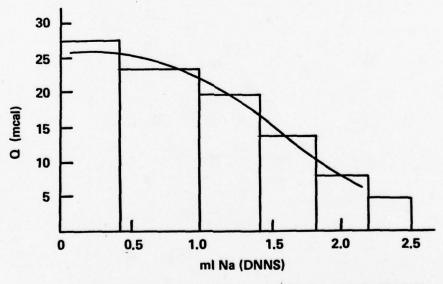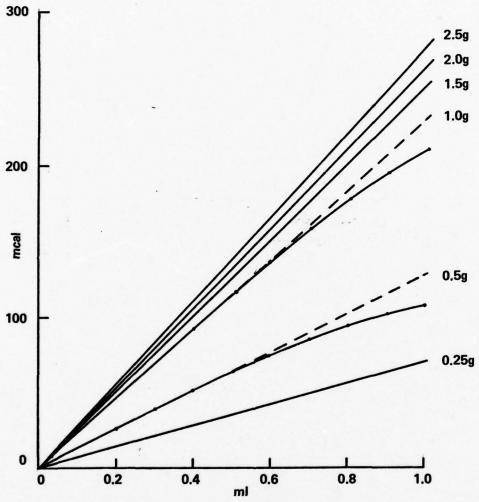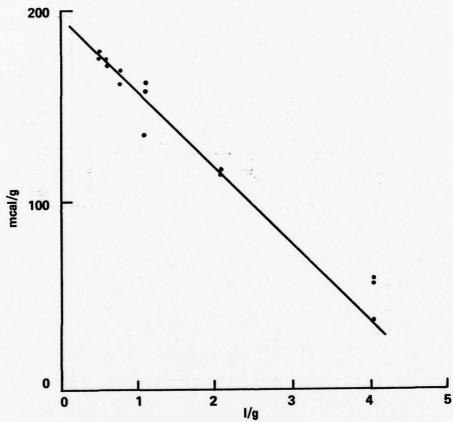

Figure IV. Thermometric Titration of Fe_2O_3 with Na (DNNS)

Figure V. Thermometric Titration Data for Various Weights
of Fe_2O_3 Titrated with 0.0321 $\underline{m}$ NaDNNS in
C_6H_{12}

Figure VI. Plot of "mcal/g" vs. "l/g" at a Surface Coverage of 19.3 μ moles/gram

RESPONSE OF SHALLOW-BURIED STRUCTURES TO BLAST LOADS

*SAMMY A. KIGER, DR.
JIMMY P. BALSARA, DR.
U. S. ARMY ENGINEER
WATERWAYS EXPERIMENT STATION
VICKSBURG, MISSISSIPPI 39180

Hardness levels of shallow-buried structures for vulnerability studies are primarily obtained from calculations using an undamped single-degree-of-freedom (SDOF) elastic-plastic model. Recent and past tests on buried structures have shown that the response has, in general, been overpredicted by both the SDOF model and by more exact analyses. In fact, no fully buried concrete structure has collapsed when exposed to simulated or actual nuclear airblast loading, with a number of structures retaining their protective capability at overpressures higher than their predicted collapse load.

Accurately predicting the response of a buried structure is complicated by the fact that the actual loading on the structure is determined in a complex manner by the response of the structure. Unlike the loads normally assumed in the analysis of surface structures, which are essentially unchanged by structural deformations, the loads acting on a buried structure are extremely sensitive to relatively small structural deformations. The differences in experimental results and calculations are attributed to insufficient knowledge of the load distribution acting on the buried structure and neglection of the influence of damping that results from structural deformation, hysteresis in soils, stress wave radiation into the soil, and shear friction at the soil-structure interface.

The Weapons Effects Laboratory at the U. S. Army Engineer Waterways Experiment Station (WES) has recently been conducting tests and analyses of shallow-buried structures to determine the influence of soil cover on the load transmitted to the structure, their static resistance and ductility, and their failure modes under static and dynamic

loading. Results of this program are being used to evaluate and modify the SDOF model for vulnerability analysis.

This paper summarizes the results from three tests conducted on reinforced-concrete box-type structures and discusses the implications of these test results on current vulnerability calculations.

Static Test Results. The objective of the static test was to experimentally determine the load-deflection behavior of the buried structure and to examine the soil-structure interaction phenomena up to the point of structural collapse. The test was conducted in the Large Blast Load Generator Facility located at the WES, Vicksburg, MS. The structure was placed in the facility and covered with sand to a depth of 1 foot. A rubber membrane was then placed over the surface of the sand, and water was pumped in to slowly apply pressure to the soil-structure system until failure occurred.

The test structure was of reinforced-concrete slab type construction where the roof, floor, and wall slabs had span-to-depth ratios of 10. The structure was 8 feet long by 2 feet high by 2 feet wide (inside dimensions). The overall wall thickness, including 0.5 inches of concrete cover on the reinforcing steel, was 2.9 inches. Principal reinforcing was 1 percent tension and 1 percent compression with approximately 1.5 percent shear reinforcement.

Figure 1 shows the load-deflection curve for the soil-structure system. The static water pressure is plotted on the vertical axis, and the roof deflection at the center of the structure is plotted on the horizontal axis. The maximum pressure sustained by the soil-structure system was about 640 psi. However, the maximum flexural capacity for the two-way roof slab assumed rigidly fixed at all edges, is a uniform pressure of 107 psi (Reference 1).

$$P_f = 107 \text{ psi} \tag{1}$$

Soil arching, defined as the ability of a soil to transfer loads from one location to another in response to a relative displacement between locations, is primarily responsible for the increase in capacity from 107 psi to 640 psi. Embedded structures that are much stiffer than their surrounding medium will tend to attract load. On the other hand, stresses will be diverted away from or around embedded structures that are less stiff than their surrounding medium. The redistribution of load is called "passive arching" when the structure is loaded above the free-field stress and "active arching" when the stresses on the structure are less than the free-field stress. An

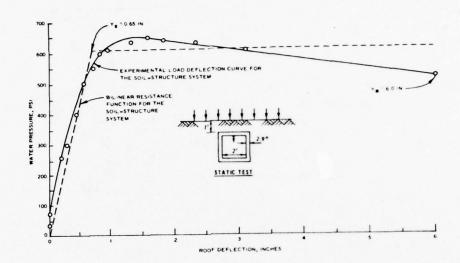excellent discussion of the soil arching phenomena can be found in
Reference 2.

Figure 1. Resistance curves for the soil-structure system.

The pressure acting on the roof and side of the structure
when the water pressure was 640 psi is shown in Figure 2. This inter-
face pressure distribution was approximated by fitting a parabola
through data taken from interface pressure gages located at each edge
of the roof, the center of the roof, the top of the sidewall, and the
center of the sidewall. Passive arching over the relatively stiff
walls of the structure has acted to increase the pressure to over
800 psi (calculated by extrapolating the parabola). However, active
arching over the flexible center of the roof has decreased the pressure
to approximately 100 psi. This shift of pressure away from the center
of the span toward the supports effectively decreases the maximum
bending moment in the structure, and can significantly increase the
resistance of the structure to the applied overpressure.

Inplane forces produced by lateral earth pressure is partly
responsible for the increased structural capacity. From Figure 2,
the thrust in the roof is 2800 lb/in. The moment-thrust interaction
diagram shown in Figure 3 was computed using a representative cross
section of the roof slab. At 2800 lb/in. of thrust, the moment
necessary to form a hinge is approximately 5800 in.-lb/in. Therefore,
the uniform pressure required to fail the roof, i.e., to form hinges
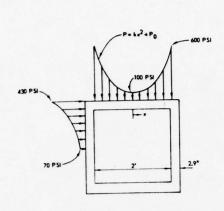
3

Figure 2. Interface pressure
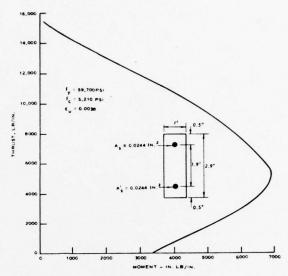distribution at
maximum surface
pressure.

Figure 3. Moment-thrust interaction
diagram.

at each support and the center of the roof, is 160 psi. Thus, the
flexural capacity of the roof including inplane forces is 160 psi.

In order to compare the computed capacity to the maximum
pressure actually sustained by the roof, an "equivalent" uniform load
will be computed based on similar moment distributions. From Figure 2,
the fixed-end moment produced by the parabolic pressure distribution on
the roof is approximately 9600 in.-lb/in. This exceeds the moment
capacity of 5800 in.-lb/in. computed from the moment-thrust interaction
diagram in Figure 3. Therefore, assume hinges have been formed at each
side of the roof and the moment at each hinge is 5800 in.-lb/in. The
moment at the center of the roof in Figure 2 necessary to maintain
equilibrium under these conditions is 7400 in.-lb/in. The "equivalent"
uniform load that will produce moments of 5800 in.-lb/in. at each end
and 7400 in.-lb/in. at the center of the roof is 183 psi. Clearly, the
inplane loading did act to increase the ultimate capacity. However,
the redistribution of the pressure on the roof due to soil arching is
primarily responsible for increasing capacity from the 107 psi in
Equation 1 to the measured 640-psi peak pressure.

Most "first-cut" design and targeting analysis procedures are
based on an undamped SDOF model using a bilinear resistance function.
The bilinear resistance for the soil-structure system as approximated

4

*KIGER & BALSARA

from the experimental load-deflection data is shown in Figure 1.
Elastic behavior is assumed up to a deflection, Y_e , and fully plastic
behavior is then assumed until failure at the maximum deflection, Y_m .
The ductility μ is defined by $\mu = Y_m/Y_e$. From Figure 1, the
ductility is about 9.

A standard calculation for the ductility involves calculating
the elastic deflection and assuming a maximum deflection. The elastic
deflection is a function of the slope of the elastic portion of the
resistance function, i.e., its stiffness, and the ultimate static
capacity, i.e., the pressure at the onset of plastic deflection. From
Biggs (Reference 3), the effective stiffness of a fixed-fixed beam of
unit width representing the roof slab is

$$K_e = \frac{307EI}{L^4} = 2650 \text{ psi/in.}$$

where I is the average moment of inertia for cracked and uncracked
sections. From Equation 1, the flexural capacity of the roof is about
107 psi. Thus, the elastic deflection, $Y_e = P_f/K_e = 0.04$ in. If a
failure deflection of 6 inches is assumed, then the ductility is 150.
Clearly, defining a "failure ductility" can be very difficult. Unfor-
tunately, most vulnerability analyses currently define failure in terms
of ductility.

In Reference 4, Cooper suggests defining failure in terms of
specific energy absorbed to failure and shows that, if e = constant
is taken as a measure of failure, then μ and T (or μ and K) are
related by

$$\mu T^2 = \text{constant}$$

or

$$\mu/K = \text{constant}$$

where T is the natural period and K is the stiffness. Therefore,
if a given ductility is defined as failure, then errors in T or K
can lead to significant errors in failure estimates.

Dynamic Test Results. Two high-explosive (HE) simulation
tests were conducted. The first experiment used 21-pound spheres of
TNT buried to midstructure depth to simulate the explosive effect of a

5000-pound general purpose bomb penetrating the soil overburden and exploding at midstructure depth. The second experiment employed a foam HEST (high-explosive simulation technique) test to simulate the overpressure from a 1-kt nuclear device at the 2000-psi overpressure range.

Both tests employed identical model structures which were twice the scale of the model used in the static test. They were reinforced-concrete slab-type models with roof, floor, and wall span-to-depth ratios of 10. The structures were 16 feet long by 4 feet high by 4 feet wide (inside dimensions). The overall wall thickness, including 0.8 inches of cover on the reinforcing steel, was 5.6 inches. Principal reinforcing steel was 1 percent tension and 1 percent compression with approximately 1.5 percent shear reinforcement. These models are approximately 1/4-scale models of a typical hardened shallow-buried structure. A more complete description of the model structure, including construction drawings, is given in Reference 5.

21-Pound HE Test Results. The model structure was covered to a depth of 2 feet for the test. Three 21-pound spherical TNT charges were used. The ch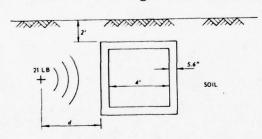arges were buried to midstructure depth and detonated at 8 feet, 6 feet, and 4 feet from the structure wall. A schematic drawing of the test configuration is shown in Figure 4. No damage resulted from the charge at 8 feet. With the charge at 6 feet, flexural cracking occurred but no measurable permanent deflection of the wall could be found. The test at 4 feet was conducted on the opposite side of the structure on the undamaged

Figure 4. Test configuration for 21-pound HE test.

wall. Structural damage resulting from the detonation at 4 feet is shown in Figure 5. Posttest measurements indicated approximately 10.5 inches of permanent deflection at the center of the wall. There were no interface pressure records obtained in this experiment; however, the pressure record measured at the soil-structure interface in a test using an identical HE charge in the same soil on a structural model with 13-inch-thick walls is given in Figure 6 and will be used in the SDOF analysis.

HEST Test Results. The foam HEST used in this study can accurately simulate the overpressure component of the airblast generated by a nuclear detonation. This test allows a large area to be
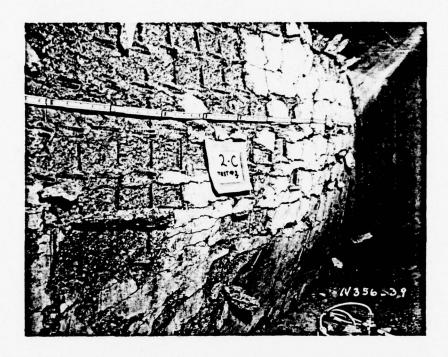
Figure 5. Structural damage from a 21-pound TNT subsurface detonation at 4 feet from the wall.
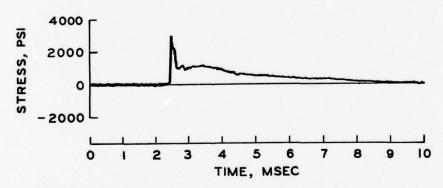

Figure 6. Soil-structure interface pressure from a 21-pound TNT subsurface detonation at 4 feet from a rigid wall.

exposed to a uniform pressure level and gives pressure durations long enough to simulate the impulse associated with a nuclear event.

The foam HEST configuration used for this test is shown in
Figure 7. The 6-inch charge cavity is filled with Styrofoam and evenly
distributed strands of detonating chord. A charge density of 0.914
pound of explosive per square foot was used to produce a peak pressure
of approximately 2000 psi, and the overburden of 32 inches of uncom-
pacted native soil was designed to confine the blast long enough to
simulate the pressure duration associated with a 1-kt nuclear device.
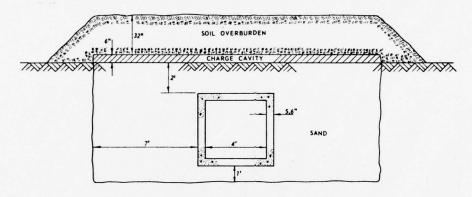

Figure 7. Foam HEST test configuration.

The airblast measured at ground level beneath the charge
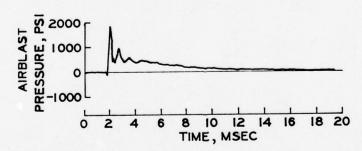cavity is shown in Figure 8.

Figure 8. Foam HEST airblast.

Interface pressures measured on the structure are shown in
Figure 9. Note that the initial peak pressure is approximately 1200
psi at each edge of the roof (Gages IF1 and IF3) and at the center of
the roof (Gage IF2). However, the pressure at the flexible center of
the roof is quickly reduced to zero and remains at zero for the dura-
tion of the airblast load. This indicates that the soil arching
phenomena so important in the static test can be equally important in

8

the case of dynamic loading. Structural damage from the HEST event was a 0.5-inch permanent deflection at the center of the roof with extensive tensile cracking of the concrete in the roof. Damage to the floor of the structure was very similar to the roof.
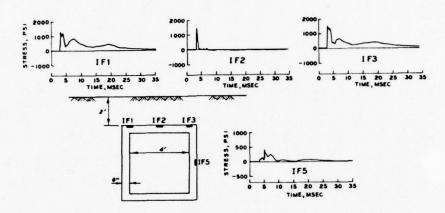

Figure 9. Foam HEST interface pressure records.

Analysis and Discussion of Results. It is assumed that the reader is familiar with the dynamic analysis of structures using a SDOF model. The essential elements of the SDOF model are an equivalent mass and resistance and a damper. Methods for calculating these parameters are given in References 1 and 3. The resistance may be either a linear or nonlinear spring and is normally approximated as a bilinear representation of the static resistance for the structure or structural element being analyzed.

21-Pound HE Test Analysis. The resistance function shown in Figure 10 was approximated from the static test results shown in
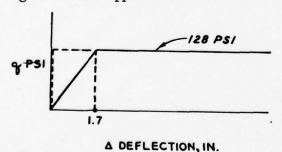
Figure 10. Resistance function for 21-pound HE test analysis.

Figure 1. The elastic deflection in Figure 1 was taken as 0.85 inch and doubled because the static test structure was a 1/2-scale model of the structure used in the 21-pound HE test. The ultimate capacity of the structure as computed using the procedures in Reference 9 is 128 psi. The 128-psi resistance includes some inplane loading but ignores soil arching and was used for the SDOF analysis

9

RIGID PLASTIC 0%

ELASTIC PLASTIC 0%

ELASTIC PLASTIC 10%

ELASTIC PLASTIC 20%

MEASURED PERMANENT DEFLECTION

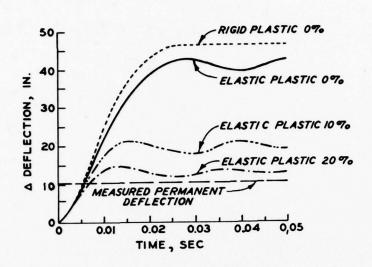Δ DEFLECTION, IN.

TIME, SEC

Figure 11. SDOF analysis results for 21-pound test.

shown in Figure 11. The percent numbers shown are the percent of critical damping in the SDOF model. Critical damping is defined as the minimum amount of damping for which no oscillations will occur under free vibration. The dynamic loading used for this analysis is shown in Figure 6. Discussion of the results in Figure 11 will be combined with the discussion of the foam HEST test analysis below.

Foam HEST Test Analysis. Since the surface loading from the foam HEST is a uniform load similar to the static load, it is assumed that the soil-structure interaction phenomena will be similar. This assumption will allow the use of the resistance function developed in the static test and shown as a dotted line in Figure 1 to be used for the SDOF analysis. Note that the displacement scale must be doubled because the static test structure was a 1/2-scale model of the HEST test structure. The loading function for this analysis is the surface airblast loading shown in Figure 8. Results of the SDOF analysis are given in Figure 12. The elastic-plastic calculation with no damping is not included in Figure 12 because it gave a 6-inch maximum deflection and could not conveniently be included on the same graph.

Discussion of Results. The results presented in Figures 11 and 12 both indicate the current vulnerability model, the undamped SDOF model, significantly overpredicts the structural response from both conventional and nuclear type blast loadings. Recent experimental results (Reference 6) and calculational results (References 7 and 8) have indicated that 20 to 25 percent of critical damping is appropriate for this structure in the buried configuration. The results in Figures 11 and 12 indicate that the 20 percent damped SDOF model can accurately predict the structural response. It is important to realize that soil arching and inplane loading were accounted for in the foam
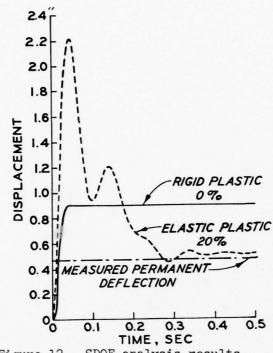
Figure 12. SDOF analysis results for foam HEST test.

HEST analysis by using the static resistance function in Figure 1 for the soil-structure system.

Figure 13 quantitatively compares the results from the experiments discussed in this report to calculations based on the most current vulnerability manual available (Reference 9).

The results presented in Figure 13 are based on calculations for a full-scale structure with inside dimensions 16 feet by 16 feet by 64 feet with walls, roof, and floor 23 inches thick (19.2 inches effective depth). The compressive strength of the concrete was 5200 psi, the yield strength of the reinforcement steel was 72,000 psi, and tension and compression reinforcement were each 1 percent. The structure was buried 8 feet deep. For all calculations represented in Figure 13, failure was taken to be 40 inches of roof deflection (4 times the deflection measured in the 21-pound HE test).

The high shear strength soil curve in Figure 13 was obtained from a 20 percent damped SDOF model assuming the same soil arching and inplane loading as would be expected in a static loading, i.e., using the resistance function in Figure 1 with the displacement scaled by 4. The zero shear strength curve represents a model with no soil arching protection provided, ultimate capacity as computed from Reference 9 at 128 psi, damping at 20 percent of critical, and elastic deflection 4 x 1.7 inches from Figure 10.

The curve labeled current vulnerability calculations in Figure 13 is based on the SDOF model recommended in Reference 9 with maximum resistance equal 128 psi. This model includes modifying the period to account for soil cover, and applying inplane forces to account for lateral earth pressure. Calculation for the roof load

includes attenuation of the overpressure with depth; however,
Reference 9 includes no damping and, for shallow depths of burial, no
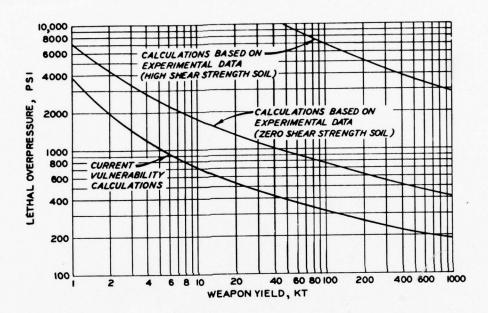soil arching.

Figure 13.   Current vulnerability calculations compared to
calculations using a damped SDOF model including
soil-structure interaction.

As an example use for the curves in Figure 13, consider the
HEST test results described above. The HEST test structure is a 1/4-
scale model of the structure calculated in Figure 13. Using cube root
scaling (charge weight scales like the cube of length), the 1-kt weapon
simulated in the HEST test scales to a 64-kt weapon in Figure 13. The
current vulnerability calculations curve indicates failure (40 inches
of roof deflection) will occur at 350 psi. The 20 percent damped SDOF
model with no soil arching indicates failure at about 900 psi. The
experimental results from the HEST test at 2000 psi produced 2 inches
of roof deflection (0.5 inch times the scale factor of 4), and the
20 percent damped SDOF model including maximum soil arching indicates
failure at the 8500-psi overpressure range from a 64-kt weapon. Based
on HEST test results, no damage would be expected at an overpressure
of 350 psi and yet the best available vulnerability analysis manual
predicts complete structural collapse at 350 psi! Clearly the vulner-
ability analysis model used for shallow-buried structures needs some

modification. Results from the test described in this report suggest that a damped SDOF model can adequately predict the response of shallow-buried structures to blast loads provided the soil-structure interaction is accounted for when computing the loads on the buried structure.

Conclusions

1. During the static test, inplane forces in the roof slab produced by lateral soil pressure increased the ultimate capacity of the structure from 107 psi to approximately 180 psi.

2. The redistribution of pressure acting on the roof of the structure caused by soil arching in the static test effectively increased the capacity of the soil-structure system to 640 psi.

3. The ductility of the structure as calculated from the load-deflection curve from the static test was approximately 9. However, the ductility is approximately 150 when calculated using the measured 6-inch maximum roof deflection from the static test and computing the elastic deflection from standard design text. Ductility is not a good parameter to use to define failure when large plastic deformations are involved.

4. Soil arching can be ignored when computing structural loads from the localized HE burst associated with conventional weapons. However, soil arching can significantly reduce the loads on a buried structure when the loading is produced by plane waves such as those associated with nuclear airblast. As indicated in Figure 13, the increased capacity due to soil arching may be as much as a factor of 10.

5. The SDOF model for calculating the dynamic response of shallow-buried structures should be damped at approximately 20 percent of critical. Then, if soil arching and inplane forces due to lateral earth pressure are adequately accounted for in computing the loads on the structure, the damped SDOF model can accurately predict the structural response from blast loads.

References

1. R. E. Crawford, C. J. Higgins, and E. H. Bultmann, "The Air Force Manual for Design and Analysis of Hardened Structures," Technical Report No. AFWL-TR-74-102, October 1974, Air Force Weapons Laboratory, Kirtland Air Force Base, NM.

2. J. W. McNulty, "An Experimental Study of Arching in Sand," Technical Report No. 1674, May 1965, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, MS.

3. J. M. Biggs, Introduction to Structural Dynamics, 1964, McGraw-Hill Book Company, New York, NY.

4. H. F. Cooper, Jr., E. J. Chapyak, and K. Y. Narasimhan, "A Parametric Study of Hard Structure Survival to Nuclear Overpressure Loading," Technical Report No. RDA-TR-104806-005, October 1977, R&D Associates, Santa Monica, CA (in draft).

5. S. A. Kiger, "Instructure Shock Environment of Buried Structures Subjected to Blast Induced Ground Shock," The Shock and Vibration Bulletin, Part 4 of Bulletin 47, September 1977.

6. R. D. Crowson and S. A. Kiger, "The Effect of Earth Cover on the Dynamic Response of Hardened Reinforced Concrete Structures," The Shock and Vibration Bulletin, Part 4 of Bulletin 47, September 1977.

7. J. E. Windham and J. O. Curtis, "Effect of Backfill Properties and Airblast Variations on the External Loads Delivered to Buried Box Structures," Technical Report, U. S. Army Engineer Waterways Experiment Station, Vicksburg, MS (in publication).

8. J. Isenberg, H. S. Levine, and S. H. Pang, "Numerical Simulation of Forced Vibration Tests on a Buried Arch," DNA Contract Report No. 4281F, 1 March 1977, prepared by Weidlinger Associates under Contract No. 001-76-C-0110.

9. J. D. Haltiwanger, W. J. Hall, and N. M. Newmark, "Approximate Methods for the Vulnerability Analysis of Structures Subjected to the Effects of Nuclear Blast," Report No. U-275-76, 15 June 1976, N. W. Newmark Consulting Engineering Services, Urbana, IL.

SILICON WAVEGUIDE LINE SCANNING ANTENNA
FOR MILLIMETER WAVES (U)

*KENNETH L. KLOHN, MR., ROBERT E. HORN, MR.
ELMER FREIBERGS, MR., HAROLD JACOBS, DR.
US ARMY ELECTRONICS TECHNOLOGY & DEVICES LABORATORY (ERADCOM)
FORT MONMOUTH, NJ 07703

INTRODUCTION

Recent demands for a very high resolution radar in terminal homing for missiles and shells and for radar surveillance in general, have generated a need for developing new concepts in low cost millimeter wave antennas. A means of providing electronic line scanning rather than mechanical scanning is desirable in order to reduce system complexity and high cost. It is especially important to eliminate the use of gimbals to mechanically scan an antenna, since they are expensive and slow. This paper describes the design and experimental findings of a novel approach for a side-looking electronic line scanner consisting of a dielectric (silicon) rectangular rod with periodic perturbations on one side. Angular scan is achieved by varying the frequency while the actual numerical values of the scan angles are a function of operating frequency, waveguide size (height and width) and perturbation spacing. An alternative approach was explored where the frequency was held fixed and the effective guide wavelength was varied electronically by modulating the conductivity of a PIN diode mounted on the dielectric waveguide. Antennas were designed for the n=-1 spacial harmonic at operating frequencies in the 55 to 100 GHz range.

DESIGN CRITERIA

The key elements in the antenna design are operating frequency (in terms of λ_o), guide wavelength, and perturbation spacing. These three values determine the nth grating lobe angle of radiation as seen in the following equation (1),(2)

1

*KLOHN, HORN, FREIBERGS and JACOBS

$$\theta_n = \sin^{-1}\left(\frac{\lambda_o}{\lambda_g} + \frac{\lambda_c}{d} n\right) \; ; \; \left|\frac{\lambda_o}{\lambda_g} + \frac{\lambda_c}{d} n\right| \le 1 \qquad (1)$$

$\theta_n = n^{th}$ beam angle from broadside (normal)
λ_o = free space wavelength
λ_g = guide wavelength
d = perturbation spacing
n = space harmonic; $0, \pm 1, \pm 2, \ldots$

The operating frequency in a system is principally chosen by application requirements such as resolution, range, and size of components. In a more practical sense, it depends on the availability of millimeter wave power sources which are small, rugged and reliable with sufficient power output for the intended application. The primary solid state sources used to date have been Gunn and IMPATT diodes. Some preliminary work has been accomplished in incorporating these devices directly in dielectric waveguides at microwave frequencies where the guides are physically much larger than those studied in this paper (3)-(5). The bulk of the experimental work reported here was done near 60 GHz where sufficient test equipment and power sources were available. Some work was completed at 70 GHz and future experimental plans are aimed at 94 GHz.

The guide wavelength, for a given material and frequency, is determined by the physical size of the guide, i.e., width "a" and height "b". Since it is advantageous from a practical point of view to avoid as much multi-signal complexity as possible, single mode operation must be maintained in the propagation of energy in the waveguide with only a single beam of energy radiating from the antenna. Thus, the designs presented in this paper which give the range (minimum and maximum) of guide sizes, allow only the E^y_{11} mode in the guide and the $n=-1$ space harmonic in the radiated energy.

Ideally, the waveguide should be kept as large as possible for any given frequency of operation as it eases fabrication problems and lessens the effect of size variations on the guide wavelength and scan angle. The maximum guide size is defined as that above which multiple mode operation is possible. The minimum guide size is defined as that below which the electric field becomes unguided. The maximum and minimum guide sizes were determined by calculating the propagation constants k_1, k_x, and k_y using the transcendental equations formulated by Marcatili (6) and solving them exactly on a programmed calculator (7) using the relationship,

$$k_2 = \sqrt{k_1^2 - (k_x^2 + k_y^2)} \qquad (2)$$

which indicated that there is no propagation down the guide (z-direction) when $(k_x^2 + k_y^2)$ exceeds k_1^2. Although Marcatili's approximate formula does not precisely match Goell's rigorous solution to the boundary value problem (8) and predict the fact that the fundamental E^y_{11} mode remains

guided no matter how small the guide's cross section, we have found experimentally that Eqn. (2) does give a good practical prediction for cut-off (9). Fig. 1 is a plot for the fundamental E^y_{11} mode and the next two higher order modes E^y_{21} and E^y_{12} for an operating frequency of 60 GHz.
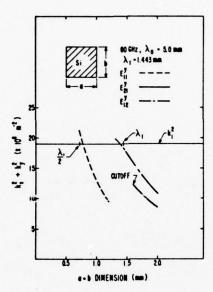

Fig. 1 Minimum and Maximum Waveguide Dimensions, 60 GHz

This plot shows that the E^y_{12} mode is the limiting mode for maximum guide size if a unity aspect ratio (a=b) is to be maintained. However, since k_x is independent of the guide height, the guide width could be made somewhat larger than the guide height; the limiting value being the cut-off point (size below which the mode will not propagate) for the E^y_{21} mode. The figure also indicates that λ_1, the wavelength in an infinite medium 1 (silicon in this case), can be used as a good first order approximation for the maximum allowed guide size and $\lambda_1/2$ as the minimum guide size. The equation for λ_1 is

$$\lambda_1 = 2\pi n_1 / \lambda_0 \qquad (3)$$

where n_1 = index of refraction of medium 1.

The figures for maximum and minimum guide size are listed in Table I. It is emphasized that these values are approximate. They do, however, provide a useful practical range of guide sizes to maintain single mode

3

*KLOHN, HORN, FREIBERGS and JACOBS

operation at the indicated millimeter wave frequencies. Later in this
paper, experimental data will be correlated with theoretical calcula-
tions for silicon waveguides well within the boundaries of the predic-
ted minimum and maximum sizes.

TABLE I

MINIMUM AND MAXIMUM WAVEGUIDE SIZES FOR FUNDAMENTAL MODE E^y_{11}

FREQUENCY (GHz)	GUIDE SIZE (mm)			
	MIN.	$\lambda_1/2$	MAX	λ_1
60	0.775	0.772	1.4	1.443
70	0.67	0.619	1.2	1.237
94	0.50	0.416	0.9	0.921

For any given waveguide size, the guide wavelength is also a function
of operating frequency. Fig. 2 shows how the guide wavelength varies
with frequency for a silicon guide, a=b=0.9 mm (maximum guide size for
single mode operation at 94 GHz). As the operating frequency falls
below the design frequency, λ_g increases rapidly. The variation in
guide wavelength when "a" and "b" are changed but kept equal to each
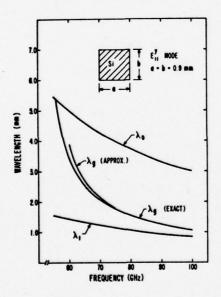other is shown in Fig. 3.
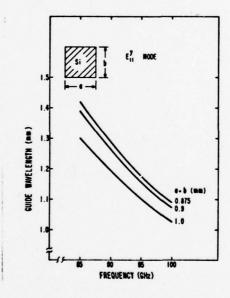

Fig. 2 Variation of Wavelength
with Frequency

Fig. 3 Variation of Guide Wavelength
with Frequency as a Function of Size

4

The smaller the guide size, the larger λ_g will be; a 10% change in guide dimensions changes λ_g approximately 5%. It will be shown later in this paper, how the changes in λ_g indicated in Figs. 2 and 3 affect the angle of radiation from periodic radiating elements. Once the physical dimensions of the dielectric guide are chosen for a given frequency, the guide wavelength will be fixed for this frequency.

Referring back to Eqn. (1), we see that for a given space harmonic (n = -1 in our case), the only remaining variable to choose is the perturbation spacing d. To obtain broadside radiation at a given frequency, d must be chosen equal to the guide wavelength at that frequency, thereby reducing θ_n to zero degrees. In actual practice, it is nearly impossible to match d= λ_g exactly, therefore, the frequency at which $\theta_n = 0°$ will be slightly different than the frequency designed to be zero.

LINE SCAN ANGLE

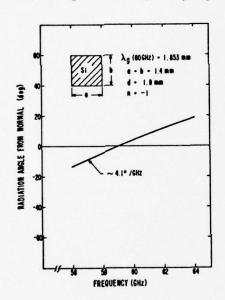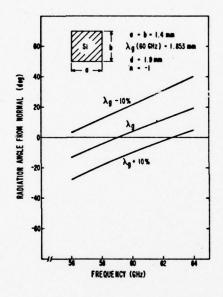Fig. 4 gives the angular scan theoretically possible (calculated from Eqn. (1)) by varying the operating frequency $f_o \pm 4$ GHz from the design frequency. A new set of propagation constants had to be calculated for each design frequency in order to determine the corresponding guide wavelength. Results of the calculations showed that the degrees of angular scan per GHz of frequency change, decreases as the design frequency increases; 4.1°/GHz at 60 GHz, 3.5°/GHz at 70 GHz, and 2.5°/GHz at 94 GHz. In addition, for frequencies above the design frequency there is the possibility of higher modes propagating in the guide. Calculations indicate that radiation angles are very sensitive to changes or errors in λ_g. Experiments conducted to measure λ_g as a function of frequency, indicated that the measured values always exceeded the calculated values (10). Fig. 5 shows the effect on the radiation angle if

Fig. 4 Radiation Angle vs Frequency (60 GHz Design Frequency)

λ_g is either $\pm$ 10% from the calculated value while the perturbation spacing is kept approximately equal to the calculated λ_g value. Over the frequency range examined, the scan angles for λ_g-10% averaged ~18° higher and for λ_g + 10%, ~14° lower than the angles calculated for λ_g.

The entire range of angular scan can be shifted more positively or negatively by changing the spacing of the perturbations as shown in Fig. 6. Closer spacing shifts the angle more negatively and conversely, farther spacing shifts the angle more positively. The slope Δ degrees/Δ frequency, remains essentially constant. For each 0.1 mm change in d, the radiation angle changes approximately 8°.
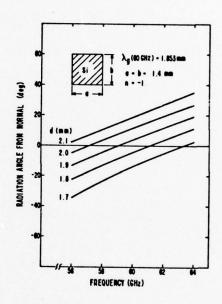
Fig. 5 Effect of λ_g on Radiation Angle

Fig. 6 Effect of Perturbation Spacing d on Radiation Angle

The range of scan angle, i.e., slope of the radiation angle curves, may be deliberately altered by changing the waveguide dimensions. As long as the dimensions stay between the minimum and maximum values, the energy will be guided and no higher modes will appear. This is illustrated in Fig. 7. Smaller guide dimensions increase the slope, thereby increasing the scan range for a given Δf_o. In each case, the perturbation spacing was adjusted to match λ_g as closely as practical. The results are tabulated in Table II.
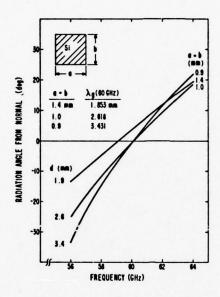
Fig. 7 Effect of Guide Size on Angular Scan with $d \approx \lambda_g$

TABLE II

ANGULAR SCAN FOR VARIOUS SIZE GUIDES (Δf_o = 8 GHz)

GUIDE SIZE a=b	PERTURBATION SPACING, d	RANGE OF ANGULAR SCAN
0.9 mm	3.4 mm	56°
1.0	2.6	43
1.4	1.9	33

A variation of guide size and/or perturbation spacing could be used to adjust the scan range. Precautions would have to be taken to avoid multi-moding which would add to the complexity of the system by producing more than one radiating beam.

EXPERIMENTAL RESULTS
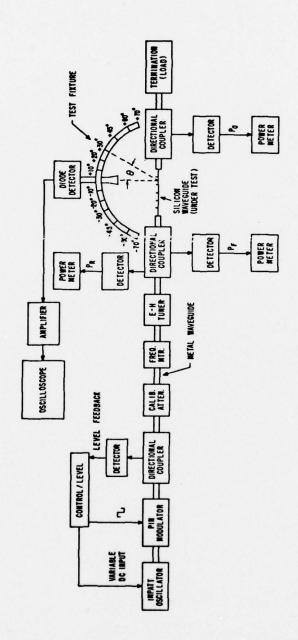
The experimental objectives were to determine: if a silicon waveguide with surface perturbations is an effective radiating structure; if the radiated beam was less than 6° in the y-z plane, where z is the direction down the guide and y is the vertical direction perpendicular to z; and if the beam is steerable as a function of frequency, or bias on a PIN diode.

Silicon waveguides were fabricated with various cross-sectional dimensions and lengths to determine their effect on the radiation angle and range of angular scan. The silicon waveguides in this paper had dimensions of 0.92 mm x 0.91 mm x 7 cm for silicon waveguide A and 0.969 mm x 1.071 mm x 10 cm for silicon waveguide B.

A series of perturbations (grating structures) on the silicon waveguide are required to provide a radiating surface. A grating structure, with grooves in a dielectric strip, for use in a leaky wave antenna has been reported by Itoh (11). The radiating structures covered in the following discussion used metal perturbations on the top of a rectangular silicon rod. The perturbations consisted of 16 to 22 rectangular copper foil stripes with dimensions of 0.3 mm x 1 mm. The stripes were cemented to the top of the silicon to provide the perturbation structure. The distance from leading edge to leading edge of each metal stripe was set to an experimentally measured silicon waveguide wavelength λ_g as measured on an unperturbed silicon waveguide and checked with calculations using the Marcatili equations. The guide wavelength was measured using a test setup as shown in Fig. 8 without perturbing stripes on the silicon waveguide and without the radiation angle test fixture. A flanged metal waveguide horn (terminated in a diode detector) was placed at a distance of less than 1 mm above the upper silicon surface in the near field of the propagated wave. The standing wave pattern on the silicon, produced by reflections due to an imperfect load, was then detected. The position between maxima points was measured using a calibrated micrometer. To obtain the average distance between maxima points, a large number of V_{max} measurements were made along the surface down the length of the guide.

The perturbations covered a length of 4 cm using 16 stripes on the silicon waveguide. The silicon waveguide was then coupled to two adjoining metal waveguide sections using the thin copper foil end plates with 1 mm x 1 mm center openings as shown in Fig. 9. The copper end plates provided support and positioned the silicon in the center of the metal waveguide. The pointed silicon waveguide extended about 1 cm into each of the two matching metal waveguide sections. In addition, the copper end plates provided a shield to prevent radiation from the metal waveguide into the air surrounding the silicon guide.

Fig. 8 Test Setup for Measurement of Radiation Angle θ vs Frequency
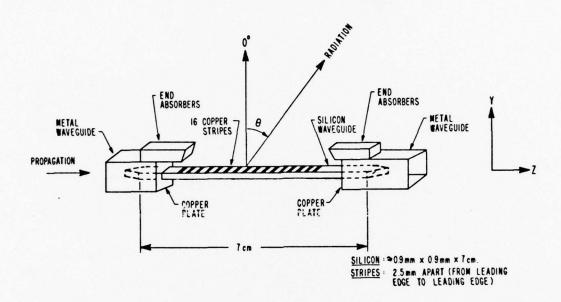
Fig. 9 Silicon Waveguide with Perturbations

With the silicon waveguide mounted in the test holder as shown in Fig. 9, the test setup (Fig. 8) was assembled and the radiation pattern from the silicon waveguide measured. The test setup was used to monitor forward power (P_f), reflected power (P_r), transmitted power (P_t), frequency (f_o), relative radiated power (P_{rad}) and radiation angle (θ). The flanged pickup horn (terminated in a diode detector and sensitive amplifier), was positioned at a distance of 20 cm from the exact center of the perturbation stripes. Most measurements were made at a 20 cm distance for mechanical convenience. However, measurements were also made at various distances in the far field (i.e., distances greater than 64 cm). The radiation patterns observed in the far field were almost identical with those observed at 20 cm. The flanged horn could then be positioned at any angle θ from $-85°$ to $+85°$ in the y-z plane on a calibrated test fixture. End absorbers (carbon compound material) were positioned directly over each end at the dielectric to metal waveguide transitions (Fig. 9) to prevent radiation leakage. The test setup used an IMPATT diode oscillator variable

in frequency from 55 to 63 GHz and a PIN diode modulator with a positive square wave modulation of about 1 kHz applied from the source. The PIN diode provided an amplitude modulated output down the waveguide and radiated out of the silicon surface. To prevent reflections from the metal waveguide to dielectric interface and thereby keep the reflected power ratio P_r/P_{in} low, the silicon waveguide was matched to the metal guide by tapering both ends. It was found by experiment that the antenna VSWR was in the order of 1.4 over a wide range of frequencies and angles of radiation. However, as the angle of radiation approached 0°, the VSWR increased to about 3.0. This rise was gradual and started to appear at about 8° from normal. The transmitted power ratio (P_t/P_{in}) through the dielectric was in the range of 10% with a slight drop in the vicinity of 0°. This behavior is to be expected from theoretical considerations when λ_g approaches the perturbation spacing d.

The fraction of power radiated for a VSWR of 1.4 can be estimated by subtracting the fraction of power reflected ($P_r/P_{in} \sim 3\%$) and power transmitted ($P_t/P_{in} \sim 10\%$). The remaining 87% of the power is either radiated or lost as dissipation in the silicon or metal stripes. The dissipation losses in silicon measured in a millimeter wave bridge were so small as to be negligible because of the high resistivity (30,000 ohm-cm) of the material. We conclude that most of the power (87%) was radiated into free space.

The radiation pattern of silicon waveguide B at a test frequency of 58.49 GHz is shown in Fig. 10a. The polar plot of relative detected power shows a negative angle $\theta = -54^o$ at the peak radiation point. The half power points occur at 50 units which indicates a beamwidth of about 4°. The theoretical value for an aperture of 4 cm is 6°. The plot was obtained by arbitrarily setting the maximum value of radiated power at 100 units on the ac amplifier meter and moving the pickup horn through an angle $\pm \theta$ on either side of the peak to obtain the detected power plot as shown.

Fig. 10b shows a typical cross-sectional radiation pattern in the plane r-ϕ which is perpendicular to the plane shown in Fig. 10a. Thus the 3-dimensional radiation pattern is narrow in the y-z plane and wide in the r-ϕ plane. Point-by-point frequency tests on silicon waveguide B yielded the relationship between radiation angle θ and frequency as shown in Fig. 11. For frequencies between 57 GHz and 65 GHz, the center of the radiated beam (θ_{max}) varied from -68° to -7°, a change of 7.5^o/GHz. The above measurements were based on operation in the fundamental waveguide mode EY_{11}. The 1 mm by 1 mm guide dimensions do not allow higher modes to propagate. The fact that these radiated beam angles are negative agrees with the theory. Since the dimensions of waveguide B are smaller than the calculated maximum for $\sim$ 60 GHz operation, λ_g, is larger than the 1.85 mm value predicted for the maximum guide size. Whenever the perturbation spacing is smaller than λ_g (1.8 mm was used for waveguide B) the effect is to shift the radiation angle more negatively as indicated in Fig. 6.
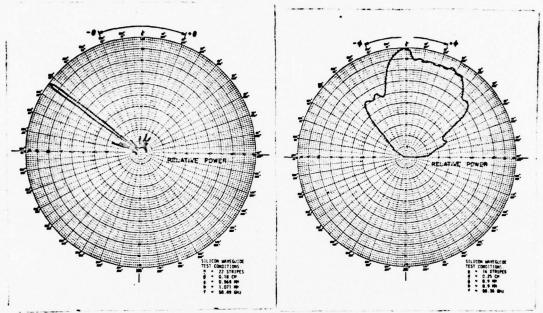
*KLOHN, HORN, FREIBERGS and JACOBS

Fig. 10a Polar Plot of Beam Angle, Y-Z Plane
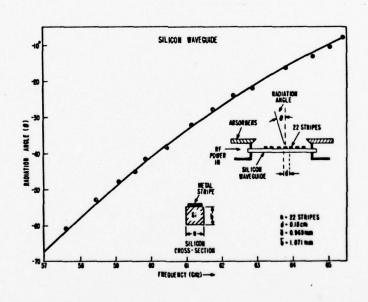
Fig. 10b Polar Plot of Beam Angle, r-φ Plane

Fig. 11 Experimental Plot of Radiation Angle vs Frequency

264

The radiation angle θ (experimental and theoretical) vs frequency for silicon waveguide B ($\bar{a}$= 0.97 mm, $\bar{b}$= 1.07 mm, $\bar{d}$= 1.8 mm) is plotted in Fig. 12. A comparison of the experimental and theoretical angles for d = 1.8 mm indicates that the experimental values (circled points) are at a more negative angle than the theoretical values for a=b=1.1 mm, but parallel, as measured from frequencies of 57 to 65 GHz. This indicates that experimental λ_g is greater than the theoretical λ_g. This agrees with previous direct measurements of λ_g using probe techniques (10).

Preliminary tests using PIN diodes to control the radiation angle have been performed. PIN diodes were cemented to the side wall of the silicon guide and biased ON in an attempt to realize a radiation angle variation as a function of PIN diode bias current. Small angular variations were observed. Emphasis will now be required on designing the optimum geometry for the PIN diode modulator.
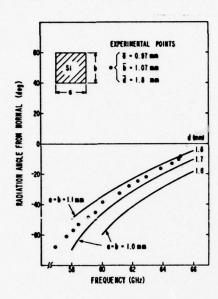
Fig. 12 Comparison of Theory and Experiment for Radiation Angles vs Frequency for Various Guide Sizes and Perturbation Spacings (Silicon Waveguide B)

SUMMARY AND CONCLUSIONS

Line scanning antennas were designed at millimeter wave frequencies (55 GHz to 100 GHz) with confirming experiments using rectangular silicon waveguides with metal stripe perturbations. Experiments indicated that a 3° - 9° beam of radiated power can be deliberately scanned through a range of angles by altering the input frequency. Using Marcatili's basic equations, the maximum and minimum waveguide sizes were determined allowing only the E^y_{11} mode for 60, 70 and 94 GHz. The key parameters of Eqn. (1), guide wavelength λ_g and perturbation spacing d, were theoretically and experimentally varied to determine their effect and criticality to the angle of radiation and range of angular scan. The theoretical results show that the radiation angle is very sensitive to changes in λ_g, a 10% change causing an angular shift of approximately 15 to 20°. The guide wavelength in turn, for a given

13

material and frequency, is determined by the physical dimensions of
the guide; width a and height b.

Small changes in the spacing of the perturbations can also
cause significant angular changes; 8° shifts occur for each 0.1 mm
change in d. These results gave an indication as to how an antenna
could be designed to cover a desired range of angular scan by altering
the physical dimensions of the guide and/or changing the perturbation
spacing.

Experiments were carried out to verify the calculated values
of λ_g by detecting the standing wave pattern and measuring the dis-
tance between maxima, $\lambda_g/2$. The experimental technique for determin-
ing λ_g indicated larger numerical values for any given frequency than
theory would predict (10), (12). However, the difference decreases
as the frequency increases. The experimentally measured radiation
angles for silicon waveguide B, plotted in Fig. 12, were compared with
theoretically calculated angles for guide sizes and perturbations
spacings closely matching the average values.

The values for a, b and d used in the theoretical calculations
bracket the actual variation in parameters for the experimental wave-
guide antennas. The results show that the observed radiation angles
fall within the theoretical limits and agreement is good, especially
since the angle θ_n is very sensitive to changes in λ_g (and therefore
a and b) and d.

The feasibility has been shown of electronically scanning through
a range of angles by varying the frequency fed into a relatively simple
antenna structure consisting of a rectangular silicon rod with copper
stripes (perturbations) attached to one surface and how the range of
scan can be altered by changing the physical dimensions of the guide
and/or the spacing between perturbations.

Investigations are continuing on fixed frequency scanners,
obtaining the change in radiation angle by modulating the effective
dielectric constant of the silicon waveguide. This is being done by
attaching PIN diodes on the side wall of the antenna. The conductivity
of the diodes is modulated by dc current injection. As the conductivity
is changed, λ_g changes and hence the angle of radiation will be varied
by electronic means.

*KLOHN, HORN, FREIBERGS and JACOBS

ACKNOWLEDGEMENT

REFERENCES

1. A. A. Oliner, Informal Communication, Class Notes from Polytechnical Institute of New York

2. A. Hessel, "General Characteristics of Traveling-Wave Antennas", Antenna Theory Part 2, Collins & Zucker, McGraw-Hill Book Company, New York, NY, 1969

3. M.M. Chrepta and H. Jacobs, "Millimeter-Wave Integrated Circuits", Microwave Journal, Vol 17, Nov 1974

4. H. J. Kuno and Y. Chang, "Millimeter-Wave Integrated Circuits", U. S. Army Electronics Command Final Report No. ECOM-73-0279-F, June 1974

5. Y. Chang and H. J. Kuno, "Millimeter-Wave Integrated Circuits", U. S. Army Electronics Command Final Report No. ECOM-74-0454-F, Oct 1975

6. E. A. J. Marcatili, "Dielectric Rectangular Waveguide and Directional Coupler for Integrated Optics", Bell System Technical Journal, Vol 48, No. 7, September 1969

7. K. L. Klohn, J. F. Armata, Jr., and M. M. Chrepta, "Transverse Propagation Constants in Dielectric Waveguides", R&D Technical Report ECOM-4242, U.S. Army Electronics Command, Fort Monmouth, NJ Aug 1974

8. J. E. Goell, "A Circular-Harmonic Computer Analysis of Rectangular Dielectric Waveguides", Bell System Technical Journal, Vol 48, No. 7, September 1969

9. H. Jacobs, unpublished. For methods used see (12)

10. H. Jacobs, G. Novick, C. M. LoCasio, and M. M. Chrepta, "Measurement of Guide Wavelength in Rectangular Dielectric Waveguide", IEEE Transactions Microwave Theory and Techniques, Vol MTT-24, No. 11, Nov 1976

11. T. Itoh, "Leaky-Wave Antenna and Band Reject Filter for Millimeter-Wave Integrated Circuits", 1977 IEEE MTT-S International Microwave Symposium Digest, Jun 21-23, 1977, pp 538-541

12. H. Jacobs, G. Novick, C. M. LoCasio, "Probe Measurements of Guide Wavelength in Rectangular Silicon Dielectric Waveguide", 1977 IEEE MTT-S International Microwave Symposium Digest, Jun 21-23, 1977, pp 118-120

MATHEMATICAL MODELS OF SKIN BURNS INDUCED BY
SIMULATED POSTCRASH FIRES AS AIDS IN THERMAL
PROTECTIVE CLOTHING DESIGN AND SELECTION

F. S. KNOX, III, T. L. WACHTEL, S. C. KNAPP
U. S. ARMY AEROMEDICAL RESEARCH LABORATORY
FORT RUCKER, ALABAMA 36362

Introduction

Fire is an ever present danger in the modern aviation environment.
Recent introduction of crashworthy fuel systems in U.S. Army helicop-
ters has dramatically lowered postcrash fire induced mortality and
morbidity (1). However, current policy states that pilots and air-
crew members will wear flight clothing which is designed to provide
some protection from the heat of such a fire. Generally, this cloth-
ing is constructed of fabrics which exhibit a high degree of thermal
stability ("nonflammable").

Proper evaluation of nonflammable fabrics requires that their
protective capability be assessed in a clinically meaningful way. A
bioassay method, using pigs as human skin analogs, was developed to
measure burn damage directly (2). This method proved useful in eval-
uating thermal protective underwear (2), four flight suit fabrics (3)
and the effect of dye deposition of skin (4). The method gives an
endpoint, burn depth, which is acceptable to clinicians and fabric
engineers alike. However, it is too costly and cumbersome for routine
fabric screening. The Thermal Analysis project has three primary
objectives: 1) Use the bioassay technique to collect a large data
base relating heat flux and exposure time to burn depth. 2) Provide
a correlation between the output of some physical heat sensors and
the burns resulting from exposure of pigs to identical fires. 3)
Develop mathematical models capable of taking the heat flux measure-
ments provided by the sensors and calculating accurately and consis-
tently the burns which would be expected.

The collection of the data base is discussed in more detail else-
where (5). The correlation between the sensors and skin has not been
established yet; although the information resides within the collected

data. This paper discusses the progress to date in meeting the third objective and a possible extension using clinical data to arrive at projected survivability.

Methods

As previously described (5), anesthetized domestic white pigs (as human skin analogs) were subjected to heat from a JP-4 fueled furnace adjusted to stimulate the heat flux, radiation and thermochemical environment of "typical" JP-4 fueled postcrash fires. Exposure times from 0.55 to 14.29 seconds and heat fluxes of 0.7 to 3.92 cal/cm^2·sec were used. Some pigs were protected by fabrics (3). In addition two sensors, a Fabric Research Labs skin simulant and an Air Force "thermoman"[1] heat sensor, were subjected to similar fires, both bare and and protected by the same standard fabrics. The resulting burns were photographed, graded using a clinical scale of 1 to 16, biopsied and graded on a micro scale of 1 to 10. Depth measurements for (a) normal epidermis, (b) normal dermis, (c) burn depth from dermal/fat border up to maximal extent of the burn, (d) dermis at burn site and (e) total depth at burn site. Corrected burn depth was calculated using the following relationship: (a+b) - c((a+b)/e).

A computerized data base was developed to manage the data from these experiments. For each burn site the following items are recorded: Pig #, Site #, smoke, template type, exposure time, heat flux furnace wall temperature, initial pig skin temperature, fabric, skin condition (natural or blackened), clinical gross grade, micro grade, epidermal thickness, dermal thickness, burn depth (epidermal/dermal border to burn), length of hair, date, time, grades from a second reading of the biopsy specimens - micro grade, normal epidermis, normal dermis, burn depth, dermal depth at burnsite, total skin depth at burnsite, corrected burn depth, computer calculated flux, computer calculated exposure time and data quality number.

In all, there are 45,752 entries for 1634 exposures from 75 pigs in the data base. The data can be retrieved via an interactive access program (PIGBOOK). Also available are other data files for furnace wall temperatures, heat fluxes, sensor responses and intraskin thermocouple responses which were recorded on FM magnetic tape and later digitized at 100 samples per second and stored on digital magnetic tape. Off-line hard copy records include ambient temperature and humidity, pig weight, sex, and data on skin cooling and water content.

The two models discussed below were programmed in FORTRAN and run on a DEC PDP 11/40 minicomputer. Preliminary development of the analytical model was carried out on an IBM 370 (6).

---

[1]"Thermoman" is an instrumented manikin developed for U.S. Air Force by Aerotherm Division of Acurex Corporation (18).

Empirical Model Development

There are many things not known about the process of burn creation in a postcrash fire. For example, details regarding heat transfer to and through fabric to skin are lacking. Thus, it is not possible to specify a priori the coupling mechanisms between the fire and skin without further detailed study. The first approach therefore was to plot the data so that some of the many possible relationships among the variables in the burn data base could be visualized. These are shown in figures 1 and 2.

TOTAL FLUX VS. GROSS BURN GRADE

GROSS BURN GRADE

TOTAL FLUX - CAL/(CM*CM)
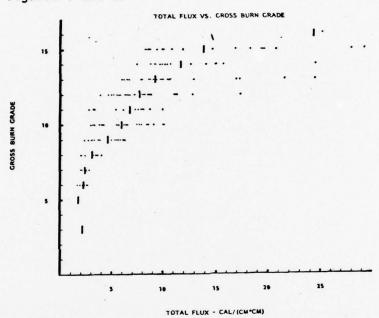
FIGURE 1

Notice the nonlinearity of the gross burn grades as they relate to total flux (cal/cm^2). Not only is this relationship nonlinear but there is a disturbing scatter to the data. The scatter is best exemplified by figure 2 in which burn depth seemed to be only generally related to total flux until the data points were identified with an exposure time. Then the data began to cluster although the scatter is still great. Identification of the processes producing this scatter is a prerequisite to satisfactory modeling. One approach to this problem is to enter the data into a general modeling system and question those observations which do not closely match the model's predictions. The aberrant observations can then be examined, any erroneous data corrected, or the observation can be deleted from the data set. The updated set of data can then be re-entered into the

3

TOTAL FLUX VS. CORRECTED BURN DEPTH
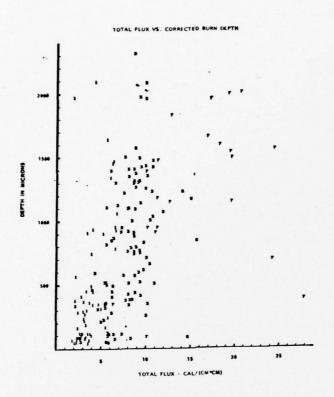
FIGURE 2

model, or another one, to further investigate relevant relationships.

This iterative procedure can be done within a framework of a multiple regression model. The problem with this procedure is that most routines available assume continuous, linear variables--an assumption which is obviously not met with the burn grade variable. This model might be suitable for analysis by burn depths, however.

A multiple discriminant model makes no assumptions about linearity of variables. It starts with groups, and develops vectors which give maximum discrimination among the different groups. These vectors then can be applied to the original observations to determine how well each matches the function's prediction as to group membership. Those observations which are not classified into their proper group may be examined for aberrant values on one or more of the predictor variables. An example of this type of process follows.

Observations with quality rankings of three or higher were selected from the data set in ascending order of gross burn grade. Predictor variables of time, flux, wall temperature, skin temperature, skin condition and time flux were extracted and calculated. This set

4

was used in a discriminant analysis with five groups: gross grades 1-4 (no dermal or epidermal burn), gross grades 5-7 (transepidermal burns), gross grades 8-10 (partial dermal burn), gross grades 11-12 (mid-thickness dermal burns), and gross grades 13-16 (full thickness dermal burns and adipose tissue burns).

Several comments on the result of the analysis are in order. The Mahalanobis D^2 can be used as an index of the degree of separation of the centroids of the groups. It is equivalent to Chi-square with number of variables times (number of groups - 1) degrees of freedom, or in this case, 24 degrees of freedom. For this analysis the value of 1261 is highly significant (see Table 1).

The analysis of classification gives, for each observation, the predicted group membership and the probability of that observation belonging to that group. In this case, with five classification groups, a random assignment would give a probability of 0.2.

Use of the classification analysis can be seen by considering observation 23 within group 1. It was predicted to belong to group 3 even though it was a group 1 observation. Looking back to the listing of the observations, it was striking that the flux given for this observation, 2.14, was notably higher than any other flux for members of this group. A subsequent review of the log book revealed that the calculated value for the flux of this observation was recorded there as 1.98 cal/cm^2. A repeated analysis using this flux value would likely yield a predicted group 2 or group 1 membership. (A mis-prediction of one group should not be considered bad since the observation may lie very close to the boundary point.) Had no predictor appeared questionable, a review of the burn grade may have indicated an erroneous grade.

This empirical model has been useful in screening the data for consistency and in assessing the importance of various predictors, e.g. the strength of flux, skin color, skin temperature and time as opposed to furnace wall temperature and total flux. But it does not have the more universal applicability of an analytical model.

Analytical Model Development

In common with previous analytical models (7-11,17) all of which were based on radiant skin burns or conductive (hot water) burns which were usually terminated at threshold blister formation (gross grade 11, micro grade 4), the USAARL/LSU Model calculates tissue temperature from heat flux and assumes that first order kinetics govern the relation between tissue temperature and damage. Although threshold blister has been a useful criterion, it does not present nearly enough data regarding the relative performance of competing fabrics.

The data base described above was collected in order to be able to expand the previous models so that they could predict burns of

TABLE I

DISCRIMINANT ANALYSIS...GROSS BURN GRADE

NUMBER OF GROUPS - 5
NUMBER OF VARIABLES - 6

Sample Sizes.. Group

Group	Size
1	25
2	34
3	61
4	121
5	143

GROUP MEANS	TIME	FLUX	WALL T.	T.°F.	SKIN T.	SKIN COLOR
1	0.887	1.012	1128.960	0.908	90.180	0.440
2	1.599	2.014	1476.177	2.509	89.601	0.413
3	3.850	2.093	1837.358	5.571	89.319	0.196
4	4.025	2.476	1734.777	8.591	91.955	0.165
5	6.085	2.709	1642.161	15.282	89.726	0.133

GENERALIZED MAHALANOBIS D-SQUARE   1261.21305

DISCRIMINANT FUNCTION 1
CONSTANT • COEFFICIENTS
-247.891 •   0.012   -61.851   0.180   1.834   5.952   16.258

DISCRIMINANT FUNCTION 2
CONSTANT • COEFFICIENTS
-354.138 •   2.130   -53.637   0.186   0.816   5.920   15.241

DISCRIMINANT FUNCTION 3
CONSTANT • COEFFICIENTS
-381.955 •   3.046   -55.353   0.106   0.680   6.074   14.372

DISCRIMINANT FUNCTION 4
CONSTANT • COEFFICIENTS
-400.905 •   2.797   -54.801   0.197   1.073   6.236   14.262

DISCRIMINANT FUNCTION 5
CONSTANT • COEFFICIENTS
-409.652 •   2.257   -55.453   0.183   1.783   6.355   14.395

EVALUATION OF CLASSIFICATION FUNCTIONS FOR EACH OBSERVATION

OBSERVATION	PROBABILITY ASSOCIATED WITH LARGEST DISCRIMINANT FUNCTION	LARGEST FUNCTION NO.
Group 1		
1	0.991	1
2	0.992	1
23	0.684	3
24	0.942	1
25	0.945	1
Group 2		
1	0.827	1
2	0.840	1
33	0.681	3
34	0.877	3
Group 3		
1	0.440	2
2	0.457	2
80	0.699	3
81	0.569	3
Group 4		
1	0.705	4
2	0.685	4
120	0.533	5
121	0.590	5
Group 5		
1	0.595	4
2	0.565	4
142	0.591	5
143	0.591	5

greater severity. The first model to be based on these data was published by Takata (11). Morse et al (18) have evaluated Takata's model in conjunction with those of Mehta and Wong (10), Henriques (7) and Stoll (9,17) using hot water burn data collected at the University of Rochester (13) as a common data base and found that Takata's model works best for dermal burns while Stoll's model works best for epidermal burns.

In plotting Takata's calculated burn depths against observed burn depths it was found that his model tends to over predict deep dermal burns caused by high heat flux and long exposure times. See figure 3.
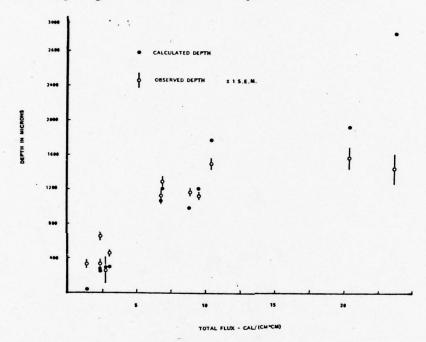
FIGURE 3
Plot of Observed Burn Depths and Calculated
Burn Depth versus Total Flux from Takata (11)

Several changes have been made to the data base since Takata's effort (5). The exposure time for each burn site was corrected to take into account differences arising from shutter system dynamics. All the biopsy specimens were reread and corrected burn depths calculated. These corrected burn depths indicate that for epidermal burns there is very slight shrinkage, followed by very slight swelling due to edema for superficial dermal burns ending with more than 40% shrinkage for mid to deep dermal burns. The heat flux measurements

7

which Takata used were hand calculated from measurements taken at 1 per second while the present data base has heat fluxes calculated from calorimeter responses digitized at 100 samples per second.

To see if an improved analytical model could be developed to explain the current data a computer program was derived as follows.

For thermal exposures of interest, skin is essentially opaque to thermal radiation and can be considered to transfer energy internally by conduction only, since exposure durations are not longer than the minimum response times reported for increased thermoregulatory system activity (14). Consequently, thermal energy transfer in skin can be described by the heat conduction or Fourier equation. In rectangular coordinates, the Fourier equation may be written as follows:

$$p \, Cp \, \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left(K \frac{\partial T}{\partial x} \right) + q \tag{1}$$

where,

p = density, gm/cm^3

Cp = heat capacity, cal/cm-sec °C

K = thermal conductivity, cal/cm-sec °C

T = Temperature, °C

x = distance, cm

q = energy source, cal/cm^3 - sec

Since skin is considered to be opaque to radiant energy, and since the source term is due only to radiant energy[1], equation (1) applies only to the surface of the skin. For all conditions in which $x > 0$, equation (1) reduces to the following:

$$p \, Cp \, \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left(K \frac{\partial T}{\partial x} \right) \tag{2}$$

Solution of equations (1) and (2) requires two boundary conditions for x, preferably at $x = 0$ and $x = L$, and initial conditions at $t=0$ for all positions $0 \leq X \leq L$. If one assumes that there is no backward flux of thermal energy at $x = 0$ (all conduction is into the skin), then the energy flux at $x = 0$ is zero and, consequently, $\partial T/\partial X = 0$. Similarly, if the problem assumes that an adiabatic backwell condition prevails at $X = L$, the fatty tissue, then the net flux out of the system at $X = L$ is 0, or $\partial T/\partial X = 0$. These two boundary conditions indicate that the system is closed and that all thermal energy added

---

[1]A simplifying assumption based on the predominance of the radiate mode of heating. May be less valid with fabrics.

8

to the system, $0 \leq X \leq L$, is distributed within the system and cannot escape.

Initial conditions are established by specifying a uniform temperature for all locations, $0 \leq X \leq L$, at time, $t = 0$.

Consequently, the system may be defined by the following mathematical model:

$$p \, Cp \, \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left(K \frac{\partial T}{\partial x} \right) + q \qquad @ \, x = 0$$

$$p \, Cp \, \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left(K \frac{\partial T}{\partial x} \right) \qquad @ \, 0 < X < L$$

(3)

$$T = To, \; 0 \leq X \leq L, \quad t = 0 \qquad \text{Initial conditions}$$

$$\frac{\partial T}{\partial x} = 0, \; x = 0, \; 0 \leq t \leq x \qquad \text{Boundary conditions 1}$$

$$\frac{\partial T}{\partial x} = 0, \; x = L, \; 0 \leq t \leq x \qquad \text{Boundary conditions 2}$$

Solution of Mathematical Model

An analytical solution to equation set (3) was not considered feasible due to the variable nature of q, Cp and K; so explicit differencing methods of numerical analysis were employed to solve the equations. Several investigators working with linear systems have found that the Crank-Nicholson six point implicit differencing method provided an excellent numerical solution (15). For the solution of equation set (3), the mathematical model, it was decided to apply the Crank-Nicholson method to the second order partial derivatives and corresponding explicit methods to the first order partials.

The implicit differencing method is noted for the characteristics of stability and convergence. Correct increment sizes yield reliable convergence. The model was implemented in FORTRAN using solution techniques of Thomas as described by Bruce (16).

This initial model was subsequently revised to allow energy flux across $x = 0$ during heating, convective heat loss at the skin surface during cooling and heat transfer into deep tissues including conduction into fat and convective cooling via the blood. The model, USAARL/LSU BRNSIM, is run interactively with most variables changeable for each run.

Since first order kinetics were assumed to apply in damaging tissue protein, tissue temperatures, T, were converted to tissue damage as follows:

$$\text{damage rate} = \frac{d\Omega}{dt} = Pe^{-\Delta E/RT};$$

$$\text{total damage} = \int_{0}^{ETIME} d\Omega/dt + \int_{ETIME}^{ITIME} d\Omega/dt$$

where
ETIME = exposure time
ITIME = total time

BRNSIM was modified to include calculations of damage rate and total damage so that it outputs damage rate, $d\Omega/dt$, for each node at each time step, total damage, Ω, for each node, and a threshold depth, where $\Omega = 1$. Threshold depth is interpolated by fitting the three Ω's nearest 1 to $Y = A + B \ln(X)$ and solving for X where $Y = 1$. Table II summarizes the predictions of this model and compares them with the observed depths. Figure 4 shows the model's calculated temperature profile at 200μ and at the fat dermal border and recorded profiles from approximately the same depths in pigs.

TABLE II

Pig	Exposure Time[***] Sec	Flux Cal/cm³·Sec	Skin Temp °C	Observed[***] Burn Depth 10-4 cm	Predicted Depth	Calculated Surface Temp °C T_{max}-T_{40sec}	Recorded[****] Temp °C T_{max}-T_{40sec}
294 LF	0.98±.01	3.31	30	257±4	283.6	99.4-39.0	49.9-45.0
294 LR	0.73±.01	3.54	31.7	222±8	252.6	76.7-38.9	69.6-45.0
294 RF	3.0	3.54	29.4	1495	*	163.5-58.4	98.2-48.4
294 RR	1.47±.01	3.92	30.6	1020±303	512.6	96.1-46.5	97.9-49.4
296 LF	3.07±.01	2.60	28.1	611±239	653.6	126.5-49.9	----
296 LR	0.99±.01	2.68	26.1	72±3	281.4	83.5-33.5	----
296 RF	8.20±.01	2.43	27.8	1488+	**	173.65-82.053	----
296 RR	1.51±.01	2.38	26.9	73±14	264.2	77.9-36.8	----

*Ω NODE 10 > 1 Interpolation scheme using nine nodes = 1155.4

+   Only one of five biopsies readable

**Ω NODE 10 > > > 1 so no depth calculation possible

***   Mean ± S.E.M.

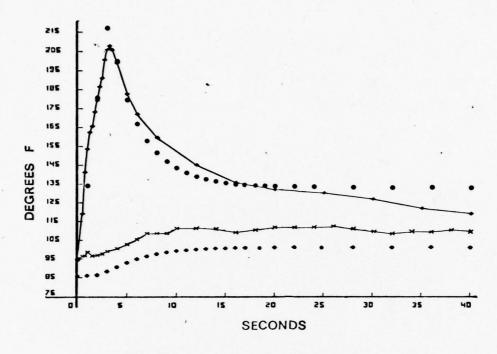****   Approximate depths of recording = 200 microns

FIGURE 4

Discussion

In the foregoing section the collection of a large data base relating heat flux and exposure time to the severity of skin burns experienced by pigs was outlined and the development of two different models of the relationship between thermal energy and burn severity was described. The multidiscriminate model was used to screen the data base for those data points which tended to lie outside of the general population. This technique has been useful in tracking down errors in coding. To date, the empirical approach has not been used to explore, in any great detail, the functional relationships which are indicated in figures 1 and 2. This part of the study has been deferred until the data base is in its final form with coding errors eliminated and extreme points flagged. Hopefully, the statistical model can be expanded and applied to the correlation of the output of the physical sensors with the skin burns resulting from exposure to the same thermal input.

The analytical model presented here is based on the work originated by Moritz and Henriques (7) as modified and extended over the years by Beuttner (8), Stoll (9,17), Mehta and Wong (10), Morse, et al

(18) and Takata (11). All of these modelers assumed that the damage in a burn results from alterations in protein structure and that these alterations proceed according to first order kinetics with a threshold at or near $44^{\circ}C$ (some authors use $45^{\circ}C$) and high energies of activation. These authors have based their models on burns induced in pigs, rats and humans by contact with hot water and by exposure to radiant sources such as carbon arc lamps. For the most part these data are limited to burns which present a threshold blister at 24 hours or less.

The only large data base available which includes more severe burns was collected at the University of Rochester in the 1950's and early 1960's using a carbon arc lamp as a source (12). These investigators were interested in studying burns which would be produced by the flash from atomic weapon detonation. Since the primary concern here was the threat of a postcrash fire, the collection of a large data base of flame induced severe burns was prerequisite to the extension of existing models or, if need be, development of new models.

The current analytical model assumes a constant temperature profile within skin, the possible incorrectness of this is revealed in the temperature profile recorded at the fat/dermal border and the computer simulation in figure 4 which shows that the initial recorded surface temperature, the one used in the simulation, was approximately $85^{\circ}F$ while the fat/dermal border measured approximately $95^{\circ}F$. In a number of such recordings a discrepancy, while not this large, seems to exist between the surface and deeper temperatures within the skin. The current version of the model does not take this into account but merely assumes that the surface temperature is distributed evenly throughout the skin depth.

The conductivity profile used in BRNSIM is that used by Morse, et al (18). This profile is based on their extensive review of the literature and is a compromise among many possible profiles including a constant thermal conductivity with depth. The product of heat capacity and density is chosen to be 1.0 throughout the skin depth except in the fat where it is 0.5. These figures were also adopted from Morse et al (18). The working values for the coefficients and exponents in the damage equation have been those of Takata (11) but because of the interactive nature of the program they can be set equal to any value including those previously reported by other authors.

As can be seen in Table II, for certain exposures, e.g. pig number 294 LF, the predictions of the model are reasonably accurate. For exposures of longer duration and high heat flux, for example pig 294 RF or 296 RF, the model fails to calculate an experimental depth because the total damage at the deepest node is greater than 1. This problem is clarified by referring back to the time temperature profile,

figure 4, in which the recorded skin temperatures are seen not to exceed the boiling temperature of water; while the computer simulation is seen to overshoot this temperature. The calculated peak surface temperature for pig 296 RF, Table II, is 173.65°C and the final temperature at 40 seconds post burn is still 81.6°C. Clearly the peak temperature is too high because this version of the model failed to take into account the water boiling for these hotter, longer exposures. Moreover, the calculated cooling phase of the tissue failed to follow the actual cooling of the tissue indicating that the heat trapped in the tissue tended to remain in the model while the heat in the real tissue was conducted deeper into the fat and/or was pulled away by the circulation.

It is clear then that several changes are required to make this preliminary version of BRNSIM conform to the physiological situation. First, an algorithm to account for tissue water boiling is required so that the tissue temperature does not exceed that of boiling water until the energy utilized in converting tissue water to steam has been accounted for. This will control peak temperature but not heat loss. Secondly, the loss of heat to deep structures and to the circulation must be adjusted. Loss of heat to the circulation is complicated by the fact that in the more severe burns the circulation is compromised by the thermal coagulation of blood components resulting in the typical picture of veno- and arteriostasis seen in clinical situations. Lastly, it will be important to express observed burn depths as corrected burn depths for the more severe burns in order to account for the extreme thermal shrinkage seen in these more severe burns. For instance, had Takata (11) used corrected depths (which were unavailable at the time) his model would not have over predicted severe burns nearly so much.

In this regard, it should be noted that the values reported for very deep burns, such as experienced in pig 296 RF where there are four missing values, are biased toward shallow burn depths by the difficulty in sectioning severely burned skin. The data from these most severe burns will require further analysis in order to determine what the appropriate depths really are.

So far the progress in formulating an analytical model which adequately predicts severe thermal injury has been encouraging and leads to speculation about possible uses of such a model in evaluating thermal protective fabrics. Fabrics exposed to a simulated post-crash fire can be evaluated for their heat transfer properties using heat sensors and the transferred heat can in turn be evaluated in terms of its potential for creating burns using the optimized version BRNSIM. The final step would be the relation of the calculated burn depths and an assumed burned area, or a measured burn area from instrumented manikins, to survivability of patients at various ages

and sexes.

As an example, consider a hypothetical case in which fabric A gives a calculated burn depth of 2000μ (full thickness burn) and fabric B a depth of 1500μ. If a constant area of burn, 30%, is assumed and the pilot is less than 34 years old, then his survivability might be 94% for fabric B but only 71% for fabric A (19). On the other hand if A were 1900μ then given the accuracy of available clinical information, there would be no difference in survivability. This example points the way toward a method of quantifying the importance of improvements in protection. But to be really useful it will require a somewhat better model and more precise clinical information regarding the relationship between burn depth and survivability. ∕

Conclusion

Progress towards supplanting the USAARL bioassay method for thermal fabric evaluation by laboratory methods involving heat sensors and a mathematical model is encouraging. Implementation will require minor changes in the model, BRNSIM, to make its output conform more closely to observed tissue temperatures and the addition of a routine to convert sensor temperatures to heat flux. Consideration of survivability will require more precise clinical data relating burn depth to clinical outcome.

REFERENCES

1. Knapp, Stanley C., Pierre Allemond. Tenth Scientific Session of JCAP, Halton, England, Sept. 6-12, 1976.

2. Knox, Francis S. III, George McCahan, Jr., Thomas L. Wachtel. Aerospace Medicine 45(8):933-938, 1974.

3. Knox, Francis S. III, Stanley C. Knapp, Thomas L. Wachtel, George R. McCahan. Tenth Scientific Session of the Joint Committee on Aviation Pathology, Halton, England, Sept. 6-12, 1976. ·

4. Knox, Francis S. III, Thomas L. Wachtel, George McCahan, Stanley C. Knapp. Joint Committee on Aviation Pathology, Halton, England, Sept. 6-12, 1976.

5. Knox, Francis S. III, Thomas L. Wachtel, Stanley C. Knapp. Proceedings AGARD-ASMP Specialists Meeting, USAARL, Ft. Rucker, Alabama, May 1-5, 1978.

6. Reneau, D.D. and Nelson O'Young. Consultation Reports to F. S. Knox, U.S. Army Contract DABT01-75-C-0257 and DAMD-17-77-C-7004, 1976, 1977, 1978.

7. Henriques, F.C. Arch. of Pathology, Vol. 43, 1947.

8. Buettner, K. J. Appl. Physiology, Vol. 5, 1952.

9. Stoll, A.M. and L.C. Green. J. Appl. Physiol., Vol. 14, 1959.

10. Mehta, A. and F. Wong. Fuels Research Laboratory, Massachusetts
    Institute of Technology, 1973.

11. Takata, A. Aerospace Med. 45(6):634-637, 1974.

12. Lyon, J.L., T.P. Davis, and H.E. Pearse. University of
    Rochester Atomic Energy Project Report UR-394, 1955.

13. Lyon, J.L., A.J. Emery, Jr., H.E. Pearse and T.P. Davis.
    University of Rochester Atomic Energy Project Report UR-401, 1955.

14. Lipkin, M. and J.D. Hardy. J. Appl. Physiol., Vol. 7, September,
    1954.

15. Crank, J. and P. Nicholson. Proc. Cambridge Phil. Soc. 43:50,
    1947.

16. Bruce, G.H., D.W. Peaceman, H.H. Rackford and J.D. Rice. Trans.
    A.I.M.E. 198:79, 1953.

17. Weaver, J.A., and A.M. Stoll. Aerospace Medicine, January 1969.

18. Morse, Howard, Glenn Fichner, Richard Brown. Aerotherm TN-75-26
    Jan. 1975.

19. Feller, Irving, Jarius D. Flora, Richard Bawol. JAMA 236:1943-
    1947, 1976.

ACKNOWLEDGEMENTS

*KOVATCH, WELLDE & HOCKMEYER

RESEARCH ON IMMUNIZATION AGAINST
AFRICAN SLEEPING SICKNESS

*ROBERT M. KOVATCH, LTC, VC
BRUCE T. WELLDE, MS
WAYNE T. HOCKMEYER, MAJ, MSC
U.S. ARMY MEDICAL RESEARCH UNIT (WRAIR) KENYA
NAIROBI DEPARTMENT OF STATE
WASHINGTON, D.C. 20520

African trypanosomiasis is an important disease of man and domestic animals on the African continent involving the area 15° North -29° South latitude (27). This Glossina (tsetse fly) infested area encompasses a land mass of approximately 4 million square miles. The disease of man, commonly referred to as African sleeping sickness, is present in endemic foci throughout this area making it hazardous or uninhabitable for humans. Insect vectors capable of transmitting the human disease, however, occur over a greater area increasing the potential for spread of this disease. Historically, large scale outbreaks occurred following the introduction of the trypanosomes into a previously unexposed population (5,15). During the period 1896-1906, over half a million people died in the Congo basin as the disease spread from the mouth of the river. During the epidemic of 1900-1910 in Uganda near Lake Victoria, an estimated 200,000 people were killed by the disease. More recently, in the 1930-1940 epidemic in East Africa, an estimated 11,500 human deaths were reported. Epidemics were followed by the establishment of persistent endemic foci. The current toll of more than 10,000 new cases reported annually, is considered to be an underestimate (3,4).

Animal trypanosomiasis is far more widely distributed. It may preclude or greatly reduce the production of domesticated animals for food or draft purposes. This deficit contributes to nutritional deficiencies and restricts social development (26).

In view of the increased awareness of our government in the African continent and the military mission to maintain a readiness

*KOVATCH, WELLDE & HOCKMEYER

capability to deploy and maintain effective military forces in any area of the world, the DOD supports research programs investigating various aspects of this problem.

There are 2 species of trypanosomes which cause human trypanosomiasis: Trypanosoma rhodesiense and Trypanosoma gambiense. Differentiation is based on the clinical syndrome as the parasites are morphologically indistinguishable (12,17). The Rhodesian form is a more acute disease which usually terminates in 1-12 months if not treated. It is characterized by recurrent fever, headache, irritability, insomnia, anemia, local edema with rapid development of congestive heart failure and often neurologic disorders. Gambian trypanosomiasis generally follows a less acute clinical course but terminates fatally when the central nervous system (CNS) becomes affected. Clinical features include marked lymph node enlargement, drowsiness, headache and ataxia. In some patients onset of CNS disease is delayed 2-6 years after demonstration of parasitemia.

A limited number of relatively toxic drugs administered by injection are available for treatment of the disease in man (1,19,20, 21, 31). Instances of drug resistance have been reported for all drugs now in use (31). Treatment with Suramin (Bayer 205) or pentamidine is fairly effective for either the Rhodesian or the Gambian form before CNS involvement has occurred. However, because CNS penetration of trypanosomes may occur early, additional follow-up treatment with drugs able to cross the blood - brain barrier is recommended. To cross this barrier, the arsenically based melaminyl compounds are now being used. Hospitalization is required for therapy because of the often severe nervous irritation and other toxic effects associated with this class of drugs (20,31).

Control of the disease has been directed toward reduction of contact between flies and man and by chemoprophylaxis. Fly numbers have been reduced by bush clearing and insecticide application. Avoidance of endemic areas has been practiced, though population movements away from such areas may be responsible for early dissimination of the human reservoir (14).

Chemoprophylaxis has been widely used in West and Central Africa with pentamidine being the drug of choice. Because of its serum binding property, pentamidine prophylaxis for the Gambian type is effective for at least 6 months (23). Suramin also has prophylactic properties for approximately 3 months but because of its nephrotoxicity and lack of effectiveness for the Gambian form, its usage has been restricted. In any case, following the introduction of trypano-

*KOVATCH, WELLDE & HOCKMEYER

somiasis into a new area it continues in an endemic form irrespective
of the control measures used (5,15).

The study of immunologic aspects of the disease which could
possibly lead to a means of control have received increasing attention
in recent years. Several workers have had limited success in immuni-
zing animals using killed suspensions of trypanosomes (13,22) or
extracts of trypanosomes as immunogens (2,28). Recent experiments have
shown that immunity can be induced in rodents, cattle and monkeys
against blood form trypanosomes by immunization with live parasites
attenuated by ionizing radiation (8,29). Experimental work with
Trypanosoma congolense and Trypanosoma brucei has not been as promising.
Attempts to immunize cattle with irradiated parasites did not result
in a sterile immunity but immunized animals mounted a humoral response
and prepatent periods were extended (6,7). A major problem with the
irradiated vaccine model is that immune animals are not protected
against heterologous strain or variant challenge (30) and no informa-
tion exists on whether these animals are protected from a natural
tsetse fly challenge.

Our studies describe immunization trials with T. congolense
in bovines in the laboratory. We also report on a field study, which
has taken place over the last 8 years, in which we determined the
antigenic similarities of T. rhodesiense occurring in an endemic area
of western Kenya.

An animal pathogen, T. congolense has been used in the
immunization trials to preclude the possibility of accidental human
infection in the laboratory or release of infected vectors into a
potential endemic area. Animal models are not unusual in the study of
human disease and we believe the immunologic similarities in terms of
antigenic variation between T. rhodesiense and T. congolense warranted
its use to preclude unnecessary risk for these studies. If anything,
our work indicates that immunization with T. rhodesiense is less
difficult than that of T. congolense (6,7,29).

MATERIALS AND METHODS

Trypanosoma rhodesiense serodemes in Lambwe Valley

To determine the antigenic composition and number of sero-
demes active in an endemic area (Lambwe Valley), isolates of T.
rhodesiense were collected from patients at the Homa Bay Hospital near
Lake Victoria, western Kenya, by members of the Kenya Medical Depart-
ment. Blood samples from the patients were then injected intraperi-

*KOVATCH, WELLDE & HOCKMEYER

toneally into rats and transported to our laboratory in Nairobi for
study. The trypanosome isolates were tested by neutralization (25)
with two separate antisera. The antisera were developed by long term
infection of two bovines, one with an isolate collected in 1972 and
the other with an isolate obtained in 1974. The duration of the
infection in these animals was 227 and 279 days, respectively. To date,
45 T. rhodesiense isolates over 8-year period have been tested.

Acquired immunity to trypanosomiasis

Animals - Bovines used in these studies were a predominantly Hereford
breed and obtained from the Veterinary Department at Kabete or from
other trypanosomiasis free areas of Kenya. All animals were dipped or
sprayed with an acaracide weekly and received periodic treatment with
Ranizole (Merck Sharpe and Dohme, B.V., Haarlem - Netherlands) to
limit helminthic infections. During periods of poor pasture supple-
mental food was provided.

Parasites - The Trans-Mara I strain of T. congolense was used in these
experiments. It was first isolated from an infected bovine in the
Trans-Mara area near the Kenya-Tanzania border in 1966 and has sub-
sequently been maintained in stabilate form at $-80^{O}C$ with occasional
passage in bovines or rodents. White albino mice obtained from the
Veterinary Laboratory rearing facility were infected with this stabi-
late. When parasitemias reached satisfactory levels the mice were
anesthetized with ether and bled by cardiac puncture. Trypanosomes
were counted in a hemocytometer and diluted with phosphate buffered
saline (pH 7.8) containing 5% glucose and 10% fetal calf serum.

Induction of immunity to blood form trypanosomes - Bovines were
infected intravenously with $1x10^4$ trypanosomes per 500lb body weight
unless otherwise noted. When animals became weak, severely anemic,
and were near death, treatment was initiated with Berenil (Farbwerke
Hoechst, Frankfurt (M) Germany) at levels of 1.05g or 2.10g active
ingredient per 660lb of body weight. This procedure of infection and
therapeutic cure was repeated until these animals were resistant to an
intravenous challenge of $1x10^4$ trypanosomes per 500lb body weight.

Tsetse fly challenge - Infected adult Glossina morsitans were used to
induce metacyclic infection. The tsetse fly colony was established
from shipments of puparia supplied by Dr. A.M. Jordan, Tsetse Research
Laboratories, Bristol, England. Standard rearing procedures were
utilized (16). Newly emerged (<24 hours post-emergence) flies were
allowed to feed on a bovine infected with the Trans-Mara I strain of

*KOVATCH, WELLDE & HOCKMEYER

T. congolense. They were exposed to the donor animal each morning for
14 consecutive days and then allowed to feed for 5 consecutive days on
each of two non-infected animals. Infection in the flies was verified
by dissecting proboscis, midguts and salivary glands from a random
sample of flies. Infected flies were allowed to feed on cattle immune
to syringe challenge with blood stage trypanosomes of T. congolense
and control animals not previously exposed to T. congolense. Each
animal received an average of 468 tsetse fly bites over a two day
period.

Monitoring of Infections - The temperatures, parasite levels and hema-
tologic parameters of experimental animals were monitored. Blood for
Giemsa stained thick and thin films were obtained from the tip of the
tail six days per week and blood for routine hematology was obtained
twice weekly from the jugular vein in disodium ethylenediamine tetra-
cetic acid (EDTA). The hematology performed included erythrocyte,
leucocyte and thrombocyte counts as well as determination of packed
cell volumes and hemoglobin concentrations. Parasitemias were esti-
mated by counting the number of trypanosomes per 100 leucocytes on
Giemsa stained thick blood films and relating these values to the
total leucocyte count per mm^3.

Verification of Immune Status - When animals which "self-cured" or
were repeatedly infected and cured with Berenil developed no detect-
able parasitemia or clinical evidence of disease after a blood induced
challenge, they were considered immune. Partially immune refers to
those animals which received a single infection and treatment. Blood
was subinoculated from immune animals into mice as an adjunct to
determining their immune status.

In conducting the research in this report, the investigators
adhered to the "Guide for the Care and Use of Laboratory Animals", as
promulgated by the Institute of Laboratory Animal Resources, National
Academy of Sciences National Research Council.

RESULTS

Survey of antigenic specificity of T. rhodesiense - During the past
eight years 45 isolates of T. rhodesiense were collected from
patients and tested against two different antisera in an attempt to
determine the extent of antigenic variation occurring in the Lambwe
Valley, western Kenya, over a period of time. As can be seen in
Table 1, antisera against two different isolates of T. rhodesiense
reacted with 60 and 61 percent of the isolates, respectively. Only 9
isolates, distributed over the 8-year period, were not neutralized by

TABLE 1

RESULTS OF NEUTRALIZATION TESTS OF ISOLATES
OF TRYPANOSOMA RHODESIENSE COLLECTED BETWEEN
1970 AND 1978 IN LAMBWE VALLEY, KENYA

Antiserum	No. Isolates tested	No. Neutralized*	Percent
Anti LVH-1	45	27	60
Anti LVH-2	44	27	61

* 9 isolates (20%) were not neutralized by either antiserum.

TABLE 5

RESULTS OF CHALLENGE BY INFECTED TSETSE FLIES OF ANIMALS
IMMUNIZED AGAINST BLOOD FORMS OF TRYPANOSOMA CONGOLENSE

Number of Animals	Group[1]	RESULT			
		Median[2] P.P. (Days)	No Detectable infection	Self Cure	Treatment
9	Immune	180	5	4	O
3	Partially Immune	14.0	O	3	O
9	Control	10.0	O	2	7

1. Immune - resistant to last blood challenge.
    Partially Immune - one infection and treatment.
2. Prepatent period.

*KOVATCH, WELLDE & HOCKMEYER

either antiserum.

Immunization Experiments - As shown in Table 2, there is a marked age resistance to primary infection with T. congolense in bovines. Eighty-five percent of 13 animals less than 1 year of age and 11% of the animals between one and two years old survived and "self-cured" the initial infection. All animals two years or older required treatment to survive. No differences in survival times were found between male and female animals of the same age group.

Five bovines, 25 to 54 weeks after "self-curing" the initial infection, received a primary challenge. All five animals were immune and showed no detectable evidence of infection while the control animals became patent and required treatment within 9 weeks (Table 3). Older animals did not "self-cure" the infection and required treatment to survive. As seen in Table 4, nine animals over one year of age that received an initial infection but required Berenil therapy, were challenged 28-128 weeks after treatment. Five of these animals "self-cured" and 4 required treatment but at times extended beyond that of the initial infection. "Self-cure" of a primary challenge seems to be dependant on a longer initial exposure to the parasite and rechallenge within a year of the last detectable parasitemia. However, there was detectable resistance even in the animals challenged over two years after the treatment of the initial infection. Five of the animals which had either "self-cured" or required treatment of the primary challenge were then given a secondary and subsequently a tertiary challenge. No detectable infections resulted in the experimental animals from either challenge while controls developed typical infections and all required treatment.

Twelve bovines, immune or partially immune to blood form trypanosomes, were subsequently challenged with T. congolense infected G. morsitans (metacyclic trypanosomes from tsetse flies). This fly-borne challenge was 8-9 months after the animal's last exposure to blood form trypanosomes. Eight control animals, not previously exposed to T. congolense, were similarly challenged. As seen in Table 5 and Fig. 1, 55% of the animals immune to blood stage trypanosomes were also completely resistant to metacyclic challenge. Though 45% became patent, they exhibited a significantly lengthened prepatent period and reduced parasitemia when compared to controls. These animals manifested no clinical evidence of disease and ultimately, all "self-cured". The partially immune animals which had received only a single exposure to blood trypanosomes, demonstrated a lengthened prepatent period and reduced parasitemia when compared to controls. They also "self-cured". Of the control group, 7 of 9 animals required treatment for survival

TABLE 2

THE EFFECT OF AGE ON TRYPANOSOMA CONGOLENSE INFECTIONS IN CATTLE

Age (Years)	Number of Animals	Median Survival* Time (Weeks-Days)	Range (Weeks-Days)	No. Self Cures (%)
0.5-1	13	>78-0	8-6 to >78-0	11 (85)
1-2	11	24-4	5-5 to >78-0	2 (11)
2-3	11	11-5	5-5 to 30-6	0 (0)
3-4	5	6-3	6-1 to 13-6	0 (0)
4-5	2	6-8	4-2 to 9-0	0 (0)
5-6	2	8-1	8-0 to 8-3	0 (0)

* Based on time to treatment or day of death.

TABLE 3

RESULTS OF PRIMARY CHALLENGE OF PREVIOUSLY INFECTED, SELF-CURED ANIMALS

	Initial Infection					Primary Challenge (1x10⁴/500lbs.)			
An. No.	Age (Yrs)	Sex[1]	Dose per 500lbs.	P.P.[2] (Days)	Last patent Parasitemia (Wks-Days)	Interval[3] (Wks-Days)	Age (Yrs)	P.P.[2] (Days)	Result
1	0.5	M	2.8×10^3	8	54-4	25-0	2.0	Not Patentable Infection	No Detectable Infection
2	1.3	MC	1.0×10^4	5	61-1	31-6	3.0	"	"
3	0.3	F	1.0×10^4	5	56-2	36-5	2.0	"	"
4	0.5	M	2.9×10^3	6	30-5	48-5	2.0	"	"
5	1.4	MC	1.0×10^4	5	38-2	54-5	3.1	"	"
Average of 3 control animals for primary challenge							3.1	4.7	T (9-3)[4]

1. F - Female; MC - Male Castrated; M - Male.
2. Prepatent period.
3. Time between last patent parasitemia and challenge.
4. Treated (Time since challenge).

TABLE 4

RESULTS OF PRIMARY CHALLENGE OF PREVIOUSLY INFECTED AND TREATED CATTLE

			Initial Infection				Primary Challenge (1x10^4/500Ibs.)		
An. No.	Age (Yrs)	Sex[1]	Dose per 500Ibs.	P.P.[2] (Days)	Time to Treatment (Wks.-Days)	Interval[3] (Wks.-Days)	Age (Yrs)	P.P. (Days)	Result (Wks.-Days)
6	1.0	F	6.8x10^6	3	7-0	28-6	1.7	14	S.C.(17-0)[4]
7	4.4	F	1.0x10^4	6	9-0	30-0	5.2	10	S.C.(11-5)
8	2.7	F	1.0x10^4	5	11-5	42-2	3.8	18	S.C.(15-5)
9	2.6	F	1.0x10^4	5	6-6	47-1	3.7	13	T.(36-6)[5]
10	1.9	F	1.3x10^5	5	28-0	71-4	3.9	14	S.C.(4-4)
11	1.6	MC	8.4x10^3	6	5-5	86-0	3.4	8	T.(21-3)
12	1.9	MC	1.0x10^4	6	5-5	86-0	3.7	6	T.(11-5)
13	2.3	MC	1.9x10^4	5	11-0	122-5	4.9	6	T.(27-0)
14	3.4	F	1.3x10^4	5	5-1	128-4	6.0	6	S.C.(29-2)

Average of 8 control animals for primary challenge   4.1   5.5   T.(9-4)

1. F - Female; MC - Male Castrated.
2. Prepatent period.
3. Time between treatment and challenge.
4. S.C. - Self Cure (Time of last patent parasitemia after challenge).
5. T - Treated (Time since challenge).

Figure 1

Average daily parasitemias of immune and control animals challenged by

Trypanosoma congolense infected tsetse flies (*Glosina morsitans*).

UNCLASSIFIED

*KOVATCH, WELLDE & HOCKMEYER

while 2 "self-cured". The 2 that "self-cured" were 1-2 years old and
reflected the age resistance to infection previously described.

DISCUSSION

Three basic problems must be resolved before immunization
could be considered a method for control of trypanosomiasis. The first
is concerned with the number and stability of serodemes persisting in
an endemic area over a period of time. Gray has shown antigenic simi-
larities in both T. brucei and T. gambiense isolates, respectively,
collected from various geographic areas of Nigeria leading one to
suspect that a limited number of serodemes were involved in each
species (9,10). In our studies, the results of neutralization tests of
T. rhodesiense isolates collected from individuals in Lambwe Valley,
over an 8-year period, showed a marked similarity of antigenic compo-
sition. Both antisera used in the tests neutralized the same as well
as different isolates which were presumably, all variants of the same
serodeme. The 9 isolates which were not neutralized by either anti-
serum may belong to the same or to one or more different serodemes.
These findings demonstrate the limited diversity and stable antigenic
composition of the parasite population. This reduces the likelihood
that erratic changes take place in the antigenic character of trypa-
nosomes in a given area and indicates that a successful vaccine could
retain its protective capacity over a relatively long period of time.

The second problem deals with antigenic variation; a process
utilized by the parasite to evade the hosts immune response (11,18,24).
Immunization by"self-cure"or infection and treatment results in expo-
sure of the host to a broad spectrum of antigenic variants. The pro-
tection induced by these procedures is apparently manifested by an
anamnestic response to most of the antigenic variants of one serodeme.
While these methods of immunization are not practical procedures for
man, it indicates the important finding that a finite number of anti-
genic variants are produced and that mutation is probably not involved
to a great degree in the process of antigenic variation. Furthermore,
it appears that exposure of a host to a significant proportion of
variants of a serodeme results in protection against the entire sero-
deme.

The third and perhaps most important problem in the develop-
ment of a vaccine involves the question of whether or not animals
immune to blood forms are protected against tsetse fly challenge.
While animals have been immunized against antigenic variants by various
methods, (2,6,7,8,13,28,29), they are resistant only to homologous
challenge and their resistance to tsetse fly challenge has not been

UNCLASSIFIED

*KOVATCH, WELLDE & HOCKMEYER

tested. In our experiments employing natural tsetse fly challenge of animals previously shown to be immune to blood forms, we found a marked degree of immunity. Most immunized animals were completely protected and showed no parasitemia or clinical evidence of trypanosomiasis. The immunized animals which became patent did so at a later time than controls and only a transient low level parasitemia developed. These transient parasitemias were not accompanied by clinical symptoms. All controls developed typical T. congolense infections characterized by progressive weight loss, anemia, leucopenia and thrombocytopenia.

In summary, we have demonstrated that the antigenic character of the parasite population of T. rhodesiense from an endemic area was composed of perhaps as few as one serodeme which was antigenically stable over an 8-year period. We also found that immunity can be induced to blood and tsetse fly (metacyclic) forms by exposure of experimental animals to a broad spectrum of antigenic variants of the same serodeme. The sterile immunity described in our studies was long lasting. We believe these findings enhance the likelihood of immunologic control of trypanosomiasis.

REFERENCES

1. APTED, F.I.C. 1960. Nitrofurazone in the treatment of sleeping sickness due to Trypanosoma rhodesiense. Trans. R. Soc. Trop. Med. Hyg., 54, 225.

2. CROSS, G.A.M. 1975. Identification, purification and properties of clone - specific glycoprotein antigens constituting the surface coat of Trypanosoma brucei. Parasitol., 71, 393.

3. de RAADT, P. 1976. African sleeping sickness today. Trans. R. Soc. Trop. Med. Hyg., 70, 114.

4. de RAADT, P. and SEED, J.R. 1977. Trypanosomes causing Disease in Man in Africa. Chapter 5, 175. In: Parasitic Protozoa Vol. 1 ED. J.P. Kreier, Academic Press, New York.

5. DUGGAN, A.J. 1970. An Historical Perspective. P. XLI In: The African Trypanosomiasis ED. H.W. Mulligan. Wiley - Interscience, New York.

6. DUXBURY, R.E., ANDERSON, J.S., WELLDE, B.T., SADUN, E.H. and MURIITHI, I.E. 1972. Trypanosoma congolense: Immunization of mice, dogs and cattle with gamma-irradiated parasites. Exp. Parasitol., 32, 527.

*KOVATCH, WELLDE & HOCKMEYER

7. DUXBURY, R.E., SADUN, E.H., WELLDE, B.T., ANDERSON, J.S. and MURIITHI, I.E. 1972. Experimental infections with African trypanosomes. IV Immunization of cattle with x-irradiated African trypanosomes. Trans. R. Soc. Trop. Med. Hyg., 66, 349.

8. DUXBURY, R.E., SADUN, E.H., ANDERSON, J.S. 1972. Experimental Infections with African trypanosomes. II Immunization of mice and monkeys with gamma-irradiated recently isolated human strain of Trypanosoma rhodesiense. Am. J. Trop. Med. Hyg., 21, 885.

9. GRAY, A.R. 1970. A study of the antigenic relationships of isolates of Trypanosoma brucei collected from a herd of cattle kept in one locality for five years. J. Gen. Micro., 62, 301.

10. GRAY, A.R. 1972. Variable agglutinogenic antigens of Trypanosoma gambiense and their distribution among isolates of the trypanosome collected in different places in Nigeria. Trans. R. Soc. Trop. Med. Hyg., 66, 263.

11. GRAY, A.R. 1975. A pattern in the development of agglutinogenic antigens of cyclically transmitted isolates of Trypanosoma gambiense. Trans. R. Soc. Trop. Med. Hyg., 69, 131.

12. HOARE, C.A. 1972. The Salivaria - subgenus trypanozoan. In: Trypanosomes of Mammals, Ch. 13, p. 476. Blackwell Scientific Publications, Oxford and Edinburgh.

13. JOHNSON, P., NEAL, R.A. and GALL, D. 1963. Protective effect of killed trypanosome vaccines with incorporated adjuvants. Nature (London), 200, 83.

14. MORRIS, K.R.S. 1960. Studies on the Epidemiology of Sleeping Sickness in East Africa III. The endemic area of Lake Edward and George in Uganda. Trans. R. Soc. Trop. Med. Hyg., 54, 212.

15. NASH, T.A.M. 1960. A review of the African Trypanosomiasis Problem. Trop. Dis. Bull., 5, 973.

16. NASH, T.A.M. and JORDON, A.M. 1970. Methods for Rearing and Maintaining Glossina in the Laboratory. Ch. 20, p. 441. In: The African Trypanosomiases ED. H.W. Mulligan. Wiley - Interscience, New York.

17. OMEROD, W.E. 1970. Pathogenesis and Pathology of Trypanosomiasis in Man. Ch. 31, p. 586. In: The African Trypanosomiases. ED. H.W. Mulligan. Wiley - Interscience, New York.

18. RITZ, H. 1914. Uber Rezidive bei experimenteller trypanosomiasis. Dtsch. Med. Wochenschr. 27, 1355.

19. ROBERTSON, D.H.H. 1961. The haemolytic effect of primaquine and nutrofurazone in cases of sleeping sickness with the haemolytic trait. Ann. Trop. Med. Parasitol., 55, 278.

20. ROBERTSON, D.H.H. 1963. The treatment of sleeping sickness (mainly due to T. rhodesiense) with melarsoprol. I. Reactions observed during treatment. Trans. R. Soc. Trop. Med. Hyg., 57, 122.

21. ROBERTSON, D.H.H. and KNIGHT, R.H. 1964. Observations on polyneuropathy and the disordered pyruvate metabolism induced by nitrofurazone in cases of sleeping sickness due to Trypanosoma rhodesiense. Acta Trop., 21, 239.

22. SCHILLING, C. 1935. Versuche zur schutzimfung gegen tsetsekrankheit. Zeitschrift für Immunitätsforschung und Experimentelle Therapie, 85, 513.

23. SCHNEIDER, S. 1963. Traetment de la trypanosomiase Africaine humaine. Bull. W. Hlth. Org., 28, 763.

24. SEED, J.R. 1963. Characterization of antigens isolated from Trypanosoma rhodesiense. J. Protozool., 10, 380.

25. SOLTYS, M.A. 1957. Immunity in trypanosomiasis. I. Neutralization reaction. Parasitol., 47, 375.

26. VAN den BERGHE, L. and LAMBRECHT, F.L. 1963. The Epidemiology and Control of Human Trypanosomiasis in Glossina morsitans Fly Belts. Am. J. Trop. Med. Hyg., 12, 129.

27. VAUCEL, M.A., WADDY, B.B., DA SILVA, M.A. DE A. and PONS, V.E. 1963. Repartition de la trypanosomiase Africaine chez l'homme et les animaux. Bull. Wld. Hlth. Org., 28, 545.

28. WEITZ, B. 1960. A soluble protective antigen of Trypanosoma brucei. Nature (London), 185, 788.

29. WELLDE, B.T., DUXBURY, R.E., SADUN, E.H., LANGBEHN, H.R., LOTZSCH, R., DEINDL, G. and WARUI, G. 1973. Experimental Infections with African Trypanosomes IV. Immunization of cattle with Gamma-irradiated Trypanosoma rhodesiense. Exp. Parasitol., 34, 62.

30. WELLDE, B.T., SCHOENBECHLER, M.J., DIGGS, C.L., LANGBEHN, H.R. and SADUN, E.H. 1975. Trypanosoma rhodesiense: Variant specificity of Immunity Induced by irradiated parasites. Exp. Parasitol., 34, 125.

31. WILLIAMSON, I. 1976. Chemotherapy of African Trypanosomiasis. Trans. R. Soc. Trop. Med. Hyg., 70, 117.

A MARKOV MODEL FOR RELIABILITY, AVAILABILITY AND MAINTAINABILITY
POLICY GUIDELINE OF AN AIRMOBILE COMBAT SYSTEM

*HAROLD Y. H. LAW, DR.
TIMOTHY D. EVANS, MR.
US ARMY AVIATION RESEARCH AND DEVELOPMENT COMMAND
ST. LOUIS, MISSOURI 63166

Introduction

The simulation of aircraft reliability, availability, and maintaina-
bility (RAM) is an extremely complex task which deals with details
of aircraft missions, scheduling, maintenance, supply, ground equip-
ment, manpower, etc. Implementation of such simulation often requires
complex models with laborious input preparation and tedious output
digestion. From the top-level decision makers' point of view, it is
helpful to gain an insight into the overall trend of significant
interactions among the aircraft RAM characteristics so as to formulate
overall policy guidelines in anticipation of future actions.

The objective of this paper is to describe a projection model which
will facilitate analysis of such interactions and to project the
aircraft availability at various stages of operation in a combat
scenario. This availability can be expressed in terms of aircraft
population at various stages and is a crucial piece of information
for decision and policy guideline. The scheduling, maintenance,
manpower, supply, etc., are governed to a large extent by availability.
The reverse is also true in order to improve availability and there-
fore mission effectiveness.

The model presented here is based on the mathematical concept of
Markov Chain Processes, supported by the real world RAM operational
sequences of an airmobile combat system. This model can be used as
a management tool which permits observations of the impact of proposed
actions prior to their implementation. Although the operation of an
airmobile combat system is addressed in this paper, this model is

1

flexible and can be adapted to other operating systems as well.

Modeling Philosophy

A model is a way of abstracting the real world so that the static and dynamic interrelationships are represented (Reference 1). With an appropriate model of a real world situation, we should then be able to predict certain outcomes or determine how the real world would behave if we implemented a particular alternative decision. One objective of model building is to identify the important variables and relationships and then translate a perception of the real world into these essential relationships and variables, and thus into a model which is tractable and, hopefully, computationally manageable. Along this modeling philosophy, we try to build a model that is simple which can be used as a management tool, and that will capture the essence of the RAM characteristics of an airmobile combat system to approximate the operations of a fleet of aircraft in a combat scenario. This model can then be used to project the movement of aircraft at various stages in the scenario, to study the interactions among the RAM characteristics, and to assess the impacts of decisions (policies) on the overall aircraft availability and RAM characteristics, so that guideline for a workable policy can be formulated.

Scenario of an Airmobile Combat System

In order to identify the essential RAM variables and relationships of an airmobile combat system, it is necessary to understand and define the scenario under which this system operates. Based on the flow of operation, the scenario can be aggregated into seven stages through which a fleet of aircraft will normally proceed. Let us assume that a fleet of aircraft starts at the ready pool stage (R) where refueling, arming, preflight preparation, etc., will be performed. This fleet of aircraft will be replenished from the reserve stage (S). From ready pool the aircraft will either remain in ready pool or go to combat mission stage (C). We define that attrition can occur only during combat due to component failure or combat damage. From the mission stage an aircraft can either remain in mission, go to attrition stage (A), come back to ready pool after the mission is accomplished, go to scheduled maintenance stage (SM), or to unscheduled maintenance stage (UM). Scheduled maintenance requirements are specified by the user according to flight time or calendar time. Unscheduled maintenance is based on component failure. From SM stage, an aircraft can either remain in that stage due to maintenance delay or other factors, or it can go to UM stage if component failure is discovered. It will return to ready pool when scheduled maintenance is completed. From the UM stage, an aircraft can either remain in that stage, or go to the not

2

operationally ready supply (NORS) stage due to waiting for parts.
Again it will go back to ready pool after maintenance is performed.
From NORS, an aircraft will either remain in NORS or return to UM
stage.

The above seven stages characterize in general terms the RAM dynamics
of a fleet of aircraft in operation.

Model Development

This chain of transition from stage to stage can be illustrated by a
directed graph (Figure 1). This graph specified the transitional

S = Reserve   UM = Unscheduled Maintenance
R = Ready Pool   NORS = Not Operational Ready Supply
C = Combat Mission   A = Attrition
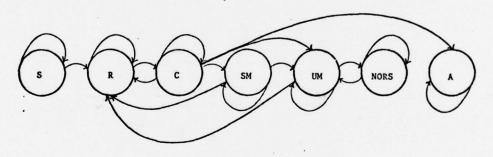SM = Scheduled Maintenance

FIGURE 1: DIRECTED GRAPH

directions of aircraft and therefore defines the movement of aircraft
and the interrelationships of the RAM characteristics. One can look
upon the stages as a set of outcomes of an experiment. The probabili-
ties of an aircraft to move from one stage to the next reflect the
RAM characteristics such as the reliability of components, the avail-
ability of manpower, equipment and parts, and the scheduling of the
resources and facilities, etc. If one can assign values of these
transitional probabilities from one stage to the other, one should
then be able to predict the outcomes. These values can be obtained
from historical data recorded in terms of the amount of time that
the system is in various stages of readiness or operability. Since
we specify a finite number of stages (seven in this case) where an
aircraft can be tracted, this set of outcomes is finite.

When considering a fleet of aircraft, we cannot say that the present condition of this fleet is independent of the past. We can say, however, that the future condition of the fleet is dependent at most on the present, i.e., it does not matter how the fleet arrives at its present condition. For example, an aircraft in mission stage depends only on its condition at the ready pool stage where preflight checking has been performed to certify the readiness of the aircraft to fly a mission. It does not depend on how it arrived at the ready pool stage. A simple model on logistic guideline of an airmobile combat system based on the concept of Regular Markov Processes was suggested by Law (reference 2). The application of such concept can be extended to the scenario described here. This scenario fits well into the concept of Markov Processes, in particular the Absorbing Markov Chain Processes.

Absorbing Markov Chain Processes

Before constructing a Markov model, one needs to make sure that the characteristics of the real world situation to be modeled satisfy the basic assumptions of a Markov process. Assuming a sequence of experiments, the outcome of each experiment is one of a finite number of possible outcomes. It is assumed that the probability of an outcome of any given experiment is not necessarily independent of the outcomes of previous experiments but depends at most upon the outcome of the immediately preceding experiment. Finally, we assume that the probability p_{ij} of an outcome on any experiment is known, given that the outcome of the preceding experiment occurred. The outcomes are called "states", and the numbers p_{ij} are called "transition probabilities". A matrix of these probabilities is called the "transition matrix". If we assume that the process begins in some particular state, then we have enough information to determine the tree measure for the process and can calculate probabilities of statements relating to the overall sequence of experiments. When these assumptions are satisfied, one can then translate the Markov processes into the operations of an airmobile combat system.

The seven stages are the states of a Markov chain. A state is called an abosrbing state if it is impossible to leave it. In the scenario of this study, the attrition stage is an absorbing state. We define that when an aircraft is attrited due to component failure or combat damage, it is lost and not salvagable and remains in the attrition stage. A Markov chain is absorbing if (1) it has at least one absorbing state, and (2) from every state it is possible to go to an absorbing state (not necessarily in one step). According to the directed graph in Figure 1, an aircraft starting at any stage will be

able to go to the attrition stage.

Model Description

The scenario under study fits well into the framework of absorbing
Markov Chain Processes, and all the assumptions for a Markov chain
are satisfied in the real world situation. Since the main objective
of this model is to tract the dynamic distribution of aircraft popu-
lation at different stages, we want the model to predict and project
the available aircraft population in any stage at any time, and to
facilitate a sensitivity analysis of the population dynamics with
respect to the RAM characteristics. This sensitivity analysis pro-
vides information for an impact study of decision on a proposed
action. The impact will lead to formulation of guideline for a
policy that may be optimal under a particular situation.

Now consider some of the questions one would like to ask and obtain
clues and solutions from a model. A policy guideline on RAM includes
considerations such as scheduling of manpower and equipment, projected
availability of aircraft, reliability of components, ordering of parts
and supplies, anticipated attrition in combat, and shipment of aircraft
from reserve, etc. Some of the questions of interest are: (1) What
is the time history of aircraft population distribution in every stage
under a prescribed policy? (2) How is this population distribution
affected by a change in policy (a change in one or more transition
probabilities)? (3) On the average, how many stages will an aircraft
go through, starting at any stages, before it is attrited, and how
sensitive is this flow towards a change in policy? (4) On the average,
how many times will an aircraft be in each stage, starting from any
stage, before it is attrited, and again how sensitive is this towards
a change of policy? (5) What is the probability that an aircraft
starting at any stage will end up in attrition?

The aircraft population at any stage is a basic piece of information
based on which the RAM policy is formulated. In a combat scenario,
one is concerned with mission effectiveness. The aircraft availability
is an important factor that contributes to mission effectiveness.
Therefore one would like to have an idea of available quantity of
aircraft so that adequate preparation can be made to keep a continuous
flow of aircraft. Since each transition probability is a description
of the decision and policy in RAM, one can conduct sensitivity
analysis and investigate how the policies affect the aircraft avail-
ability and how one can improve decision policies to provide timely
maintenance and to maintain mission effectiveness. This sets the
tone of an overall policy guideline. On a more detailed level of

management, question (3) leads to some insight on how one should prepare for the availability and scheduling of manpower, equipment, supplies, etc., at various stages to anticipate and service the flow of aircraft. Question (4) is similiar to question (3) except it is more concerned with the local scheduling and supply of a particular stage.

According to the structure of this absorbing Markov model, an aircraft will ultimately end up at the attrition stage. This is true in reality when an aircraft has accumulated sufficient number of flight hours. In the scenario under study, there is only one attrition stage. This means the probability of attrition for any aircraft will be 1. This is true but not very interesting. However, if there is more than one attrition stage in the model, i.e., attrition due to combat damage, component failure, accidents, and other causes, then the model can provide information on the probability of attrition due to various causes. This information may be of interest to the decision maker, and the Markov model is flexible to accommodate this feature. However this feature is not included in this study. The questions above are by no means all one wants to ask in formulating policy guideline but are questions of major interest and concern.

The condition of a system can be expressed as a state vector containing w states, $S(s_1, s_2, \ldots, s_w)$. If the system is in state s_i, the probability that it will be in state s_j is p_{ij} such that $s_j = p_{ij} s_i$. This transition probability can be considered as a description of the decision that takes the system from state s_i to state s_j. This decision is the result of actions implemented according to a policy. Thus the transition of a system at any state can be described by the transition matrix

$$p = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1w} \\ p_{21} & p_{22} & \cdots & p_{2w} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ p_{w1} & p_{w2} & \cdots & p_{ww} \end{pmatrix} \quad \text{such that } S_{k+1} = PS_k \quad (1)$$

where k indicates the time period and

$$\sum_{j=1}^{w} p_{ij} = 1, \ i = 1, 2, \ldots, w \text{ and } p_{ij} \leq 1.$$

For an arbitrary absorbing Markov chain, the absorbing states can be grouped together and the transition matrix can be rearranged in the following canonical form

$$
P = \begin{array}{c} \\ r \\ t \end{array}
\begin{array}{cc} r & t \\ \left[\begin{array}{c|c} I & 0 \\ \hline R & Q \end{array} \right] \end{array}
\tag{2}
$$

where I is an rxr identity matrix, 0 is an rxt zero matrix, R and Q are the partitions of the remaining elements in matrix P, and the elements in R and Q are less than unity by definition.

A theorem (reference 3) shows that the inverse matrix, $(I-Q)^{-1}$, exists and is called the fundamental matrix.

Let $N = (I-Q)^{-1}$ and $N = (n_{ij})$, $i,j = 1,2,\ldots,t$

Futhermore, a theorem (reference 3) indicates that the matrix of probability of absorption is a txr matrix B such that $B = NR$. These matrices contain important information and interpretation which are of interest to policy guideline in this study.

The elements n_{ij} of the fundamental matrix N is the expected number of times before attrition that an aircraft will be in state j if it starts in state i. This information provides an answer to question (4). The sum of n_{ij} along the row elements represents the expected number of times that an aircraft will be in a nonabsorbing state (nonattrition stage) if it starts at state i. This provides an answer to question (3). To answer question (5), we notice that the interpretation of element b_{ij} of matrix B is precisely the probability that an aircraft will be attrited at the absorbing state j if it starts at state i. We shall demonstrate below how this model can be applied to the scenario of an airmobile combat system. According to the scenario illustrated in Figure 1, the canonical form of the transition matrix can be presented as

$$
P = \begin{array}{c}
 \\ A \\ S \\ R \\ C \\ SM \\ UM \\ NORS
\end{array}
\begin{array}{c}
\begin{array}{ccccccc} A & S & R & C & SM & UM & NORS \end{array} \\
\left[\begin{array}{c|cccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & P_{22} & P_{23} & 0 & 0 & 0 & 0 \\
0 & 0 & P_{33} & P_{34} & 0 & 0 & 0 \\
P_{41} & 0 & P_{43} & P_{44} & P_{45} & P_{46} & 0 \\
0 & 0 & P_{53} & 0 & P_{55} & P_{56} & 0 \\
0 & 0 & P_{63} & 0 & 0 & P_{66} & P_{67} \\
0 & 0 & 0 & 0 & 0 & P_{76} & P_{77}
\end{array} \right]
\end{array}
= \left[\begin{array}{c|c} I & 0 \\ \hline R & Q \end{array} \right]
\tag{3}
$$

where $0 \leq p_{ij} \leq 1$

Let $A = (I-Q)$, then $N = A^{-1}$.

The time history of aircraft population distribution and the effect of change of policy on available population can be obtained by the recurrsive equation (1), $S_{k+1} = PS_k$. The sensitivity mentioned in questions (3) and (4) can be investigated through the partial derivatives $\frac{\delta b_{ik}}{\delta p_{pq}}$ or $\frac{\delta b_{ik}}{\delta a_{pq}}$ and $\frac{\delta n_{ij}}{\delta p_{pq}}$ or $\frac{\delta n_{ij}}{\delta a_{pq}}$. A moment of reflection on Cramer's rule indicates that a_{pq} can be isolated in the expressions of b_{ik} and n_{ij}. Therefore analytical expressions of the partial derivatives as a function of a_{pq} can be obtained. Let the minor of matrix A be M_{ij} and the cofactor be C_{ij} such that $C_{ij} = (-1)^{i+j} M_{ij}$. The elements of N can be expressed as

$$n_{ij} = \frac{(-1)^{j+i} |M_{ji}|}{|A|} = \frac{(-1)^{j+i}(a_{pq} C'_{pq} + K_1)}{a_{pq} C_{pq} + K_2} \qquad (4)$$

$$q \neq i, \quad p \neq j$$

where $|M_{ji}| = \sum_{\substack{h=1 \\ h \neq i}}^{t} a_{mh} C'_{mh}$, $m \neq j$ and C'_{mh} is the cofactor of M_{ji} and

K_1, K_2 are constant terms not involving a_{pq}. Therefore

$$\frac{\delta n_{ij}}{\delta a_{pq}} = \frac{(-1)^{j+i}(C'_{pq} K_2 - C_{pq} K_1)}{(a_{pq} C_{pq} + K_2)^2}, \quad p \neq j, \ q \neq i. \qquad (5)$$

Similarly,

$$\frac{\delta \sum_{j=1}^{t} n_{ij}}{\delta a_{pq}} = \sum_{j=1}^{t} \frac{(-1)^{j+i}(C'_{pq} K_2 - C_{pq} K_1)}{(a_{pq} C_{pq} + K_2)^2}. \qquad (6)$$

*LAW & EVANS

Since B = NR, where R = (r_{jk}), then

$$b_{ik} = \sum_{j=1}^{t} n_{ij} r_{jk}, \quad i,j = 1,\ldots,t, \text{ and}$$

$$\frac{\delta b_{ik}}{\delta a_{pq}} = \sum_{j=1}^{t} \frac{n_{ij}}{a_{pq}} r_{jk}, \quad k = 1,\ldots,r. \tag{7}$$

Hence the sensitivity of the flow of this system towards a policy at any stage can be obtained in equations (5), (6), and (7). The answers to the above five questions can also be obtained through a simple algorithm (Figure 2). This algorithm can be easily adapted to accommodate as many stages as one wishes to have. A listing of the algorithm is available upon request.

Model Application and Discussion

Now let us apply the model to a real system. The data in Table 1 gives the required parameters and the average condition of a helicopter fleet in 1971. Some of the results of this application are presented in Figures 3 to 7. These figures show the trend of population dynamic which is the primary information for policy making. The aircraft population distributions at all seven stages are presented in Figure 3. The aircraft availability at the stages of ready pool, mission, scheduled and unscheduled maintenance fluctuates at the early time period, and gradually stablizes. For a limited population of aircraft

FIGURE 2: FLOW DIAGRAM OF ALGORITHM

at the reserve stage, it is seen that the aircraft at that stage
diminishes as timé goes on, and the attrition accumulates at a fairly
constant rate. If a policy of skipping scheduled maintenance is in
effect, it is observed in Figure 3 that more aircraft are available
for mission, and yet more aircraft are also attrited. However, the
SM, UM, and NORS stages are not affected significantly by this policy.
When the rate from reserve to ready pool is increased as in Figure 4,
the available aircraft at ready and mission stages increases sharply
in the early time period, and attrition also increases sharply.
Therefore more manpower, equipment and facilities are needed at
the maintenance stages. It is interesting to notice that at a
later time period, aircraft is actually less available for mission
because of limited reserve and high attrition. Therefore the
reserve pool would need to be built up in order to maintain a
certain level of aircraft availability. When the rate of attrition
is increased, the drastic effects are observed in Figure 5, and the
level of availability becomes very low. It was observed that
shortening the waiting for supply at the NORS stage does not notice-
ably affect the availability. The graph of which is not shown due
to lack of space. One example of the sensitivity of the flow of
aircraft towards a change of policy is presented in Figures 6 and 7.
In this example, the sensitivity increases as the transition prob-
ability increases. In other situations, the reverse trend may be
true.

Conclusion

In this paper we have demonstrated the application of absorbing
Markov Chain Processes to analysis of reliability, availability, and
maintainability policies. A model was presented and applied to a
real world situation. This model is simple and flexible. Because
of the simplicity of the model structure, the algorithm of which
can be easily programmed and be made interactive. The Markov Chain
Processes presented here show great promise in analyzing RAM policies.
The model presented here is not intended to replace simulation models,
but can be used to gain insight into trends which would result from
overall policy changes.

References

1. Gass, Saul, Roger L. Sission, A Guide to Models in Governmental
    Planning and Operations, PB-245-269, August, 1974.

2. Law, Harold Y. H., "A Macroscopic Modeling Concept for Logistic
    Policy Guideline of an Airmobile Combat System", Proceedings of

the 15th Army Operations Research Symposium, 26-29 Oct. 1976,
Ft. Lee, Virginia.

3. Feller, William, An Introduction to Probability Theory and Its Applications, Volume I, 3rd edition, John Wiley, New York, 1970.

Table 1

Condition and Parameters of a Helicopter Fleet

Average inventory: 2,526 aircraft
Total flight hours: 1,096,510
Total flight: 2,917,955
Mean time between scheduled maintenance: 2.749 hours
Mean time between unscheduled maintenance: 2.652 hours
Mean time between all maintenance: 1.35 hours
Mean elapsed scheduled maintenance time: 5.40 hours
Mean elapsed unscheduled maintenance time: 5.195 hours
Mean elapsed maintenance time: 5.30 hours
Clock hours: 8,760

Rate of attrition, $r_1 = 0.1$ per clock hour
Rate from reserve to ready pool, $r_2 = 0.1$
Rate of launch, $r_3 = 0.1319$
Rate of leaving mission, $r_4 = 2.6611$
Rate from mission to scheduled maintenance, $r_5 = 0.0180$
Rate from mission to unscheduled maint., $r_6 = 0.0187$
Rate from scheduled to unsched. maint., $r_7 = 0.0$
Average length of mission, $t_c = 1.0$ hour

$P_{23} = 1.0 - \exp(-r_2 t_c)$
$P_{33} = \exp(-r_3 t_c)$
$P_{41} = 1.0 - \exp(-r_1 t_c)$
$P_{44} = \exp(-r_4 t_c)$
$P_{45} = 1.0 - \exp(-r_5 t_c)$

$P_{46} = 1.0 - \exp(-r_6 t_c)$
$P_{56} = 1.0 - \exp(-r_7 t_c)$
$P_{55} = r_8 = 0.1852$
$P_{66} = 0.2003$
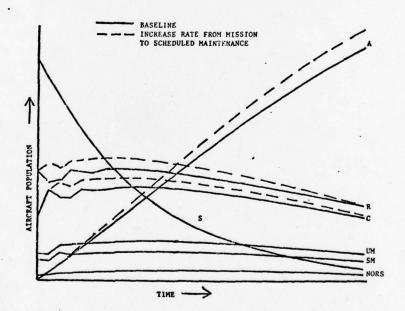$P_{67} = 0.1$
$P_{77} = 0.3165$

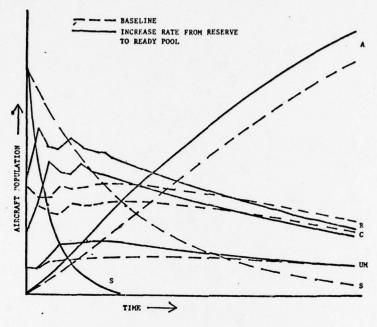FIGURE 3: DYNAMICS OF AIRCRAFT POPULATION DISTRIBUTION

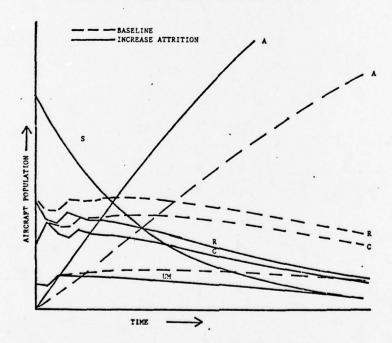FIGURE 4: DYNAMICS OF AIRCRAFT POPULATION DISTRIBUTION

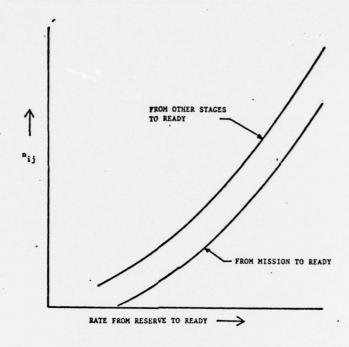FIGURE 5: DYNAMICS OF AIRCRAFT POPULATION DISTRIBUTION
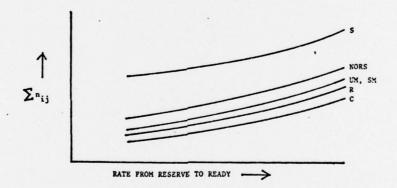
FIGURE 6: AIRCRAFT FLOW SENSITIVITY

FIGURE 7: AIRCRAFT FLOW SENSITIVITY

ENVIRONMENTAL EFFECTS ON THE MECHANICAL PROPERTIES OF
GLASS FIBER/EPOXY RESIN COMPOSITES
Effect of Static Immersion in Water on the
Tensile Strength of Crossply Laminates

*BYUNG LIP LEE, ScD, ROBERT W. LEWIS, PhD, AND ROBERT E. SACHER, PhD
ARMY MATERIALS AND MECHANICS RESEARCH CENTER
WATERTOWN, MASSACHUSETTS 02172

INTRODUCTION

The use of glass fiber-reinforced resin composites in primary
structures for Army applications has been on the increase in recent
years. They are being, or will be, employed in load-bearing areas in
several rotor blade systems - the CH-47 modification, the Advanced
Attack Helicopter, the Blackhawk, and the AH-1Q Cobra Improved Main
Rotor Blade - as well as being used in critical areas in the VIPER
system. Despite this projected usage, the questions on the stability
of the material properties upon exposure to outdoor environments re-
main largely unanswered (1,2). The outdoor weathering conditions are
extremely complex and include such factors as temperature, electromag-
netic radiation (for example, sunlight), and moisture. These are con-
sidered to be the major elements responsible for the degradation of
materials in the atmosphere (2,3).

This study is directed at determining how and why the mechan-
ical properties of fiber-reinforced resin composites are affected by
exposure to each of the above-mentioned elements with different dura-
tion, intensity, and frequency. In this phase of the study, the effect
of moisture at ambient temperature without radiation is being investi-
gated. Since precipitation in the form of rain or dew is the most
common source of moisture in the outdoor environment, immersion in
distilled water was chosen as the first exposure condition to be studied.
Among a wide variety of mechanical properties, the uniaxial tensile
strength of the laminate under constant-rate deformation was chosen
mainly because of availability and simplicity of the testing technique.

*LEE, LEWIS, and SACHER

The techniques and results developed in this work will be eventually translated to the actual systems under consideration for helicopter rotor blades.

A typical composite laminate may be built up with the fibers of each lamina oriented in a different direction. When a laminate contains laminae with different fiber orientation and is subjected to uniaxial tension, individual laminae will fail successively with increasing order of failure strain in the given direction. In a cross-ply laminate, the lamina with its reinforcing fibers aligned perpendicularly to the load direction, i.e., the 90° lamina, will fail first, thus reducing its share of load. The burden of load transfer will be increasing in the unbroken portion of the laminate (4). As a consequence, the "knee" or "yield" point appears in the stress-strain curve due to the reduction of laminate modulus (Figure 1a). Although the term "yield point" will be used hereafter based on the analogy to that in metals, underlying deformation mechanisms are entirely different. What happens in the composite laminate at this point is micro-failure followed by the formation of translaminar crack, not plastic flow or
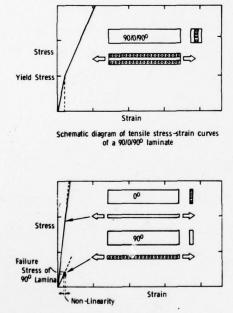
Figure 1a. Schematic diagram of tensile stress-strain curves of a 0° lamina, a 90° lamina and a 90/0/90° laminate.
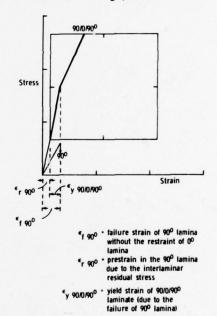
Figure lb. Approximate relationship between the yield strain of a 90/0/90° laminate and the failure strain of a 90° lamina (without the restraint of a 0° lamina).

dislocation motion. The yield point is usually far below the cata-
strophic failure point (which determines ultimate strength or strain)
of the laminate. But it is very important in the design concept,
since the "design ultimate stress" is related to the maximum laminate
stress attainable without the rupture of any lamina (5).

The yield stress or strain is also important from the view-
point of resistance to water absorption, because the occurrence of
cracks for whatever reason will allow water absorption of composites
by capillary or conduction mechanisms. Therefore, the absorption of
water into the composites can be accelerated at the yield point, pos-
sibly resulting in further deterioration of the properties. Consider-
ing these facts, the effect of water absorption on the tensile strength
of crossply (90/0/90°) laminates will be examined with particular
emphasis on the yield point.

EXPERIMENTAL PROCEDURE

Three types of glass fiber/epoxy resin composites have been
used in this phase of the study. In this report, discussion will be
restricted to the data based on E glass/Scotchply 1009 resin system.
The E glass/1009 resin laminates were prepared by a commercial manu-
facturer (3M, St. Paul, MN). The following conditions were used in
their curing and postcuring process: cure at 163 C (325 F) for 45 min-
utes under pressure of 50 psi, postcure at 177 C (350 F) for 4 hours
under contact pressure. The laminate supplied consists of three cross-
plied laminae (lamina thickness ∿0.016 inch). The laminate has a den-
sity of 2.01 ± 0.02 gram/cc and 74.9 ± 1.2 weight percent fiber (6,7).

Three sets of specimens (34 in. × 6-1/2 in. × thickness) were
used in the water absorption experiments:

(1) three-ply tension test specimens with the direction of
fiber on the outermost lamina perpendicular to the specimen axis
(90/0/90°),

(2) single-ply tension test specimens (0° or 90°), and

(3) warped two-ply specimens (0/90°).

The following procedure was used for measuring the change in
the tensile properties of 90/0/90° laminate by water absorption.

(1) The 90/0/90° specimens were cut from the plate and com-
pletely dried at 50 C under vacuum.

(2) The dried specimens were immersed in distilled water at
23 C for various periods of time.

(3) After the moisture was removed from the specimen surface by soft paper tissue, the weight gain of each specimen was measured.

(4) The specimens having a gage length of 4.5 in. were tested in an Instron tester at a crosshead speed of 0.05 in./min.

(5) Yield stress and strain were defined by the position of the knee in the stress-strain curve and ultimate strength by the catastrophic failure point. The tangent modulus of elasticity was estimated from the slope of the stress-strain curve.

The 90°, 0°, and 0/90° specimens were prepared by removing unnecessary lamina(e) from the 90/0/90° specimens which had been immersed in water for various periods of time. Since the removal of unnecessary lamina was performed by a sharp razor blade within 10 minutes, its effect on the water retention of the specimens was assumed negligible. The 90° and 0° specimens were used in the evaluation of strength and modulus of each lamina. The curvature of 0/90° specimens was estimated from the values of height and span, assuming that the arc of warped specimens was circular. All measured properties were plotted against the weight gain due to water absorption.

RESULTS AND DISCUSSIONS

The weight gain of 90/0/90° specimens due to water absorption is shown as a function of immersion time in Figure 2. It clearly shows that equilibrium has not been reached at 2500 hours of water immersion at ambient temperature.
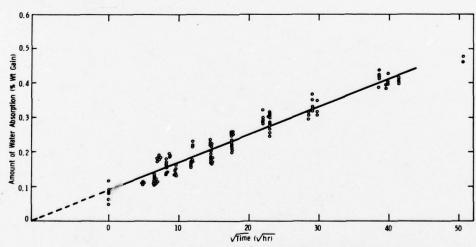

Figure 2. Amount of water absorption of a 90/0/90° laminate versus the immersion time at room temperature.

*LEE, LEWIS, and SACHER

The results of tension tests show that, at room temperature, water absorption affected the level of both yield stress and ultimate strength of 90/0/90° laminates to a considerable degree. Ultimate strength decreased continuously with an increasing amount of water absorption (Figure 3). Yield stress increased initially and then decreased, thus to form a broad maximum around the point of 0.2% weight gain (Figure 4). On the other hand, no significant change was observed in the values of moduli before and after yielding due to water absorption (Figure 5).

As shown earlier, the yield stress is very important from the viewpoint of both the design limit of the material and the resistance
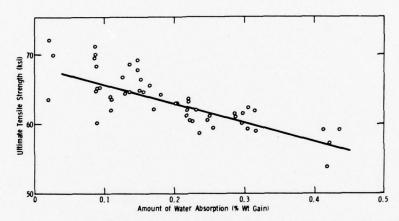

Figure 3. Effect of water absorption on the ultimate tensile strength of a 90/0/90° laminate.
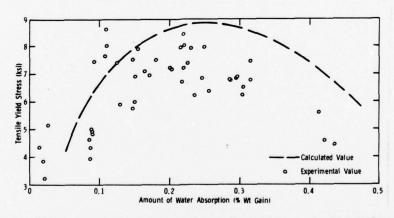

Figure 4. Effect of water absorption on the tensile yield stress of a 90/0/90° laminate.
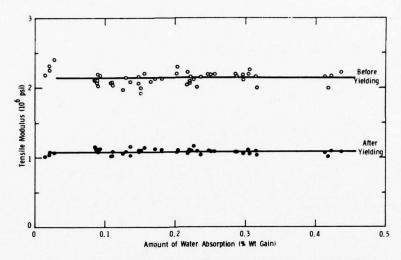
Figure 5. Effect of water absorption on the tensile modulus of a 90/0/90° laminate.

to water absorption. Considering this, discussion will be directed to the effect of water absorption on the yield stress. Since the yield stress is closely associated with the failure of the 90° lamina, discussion will start with the description of the deformation mechanism of 90/0/90° laminates in uniaxial tension.

Compared to other angle-ply laminates (8), 90/0/90° laminate exhibits a relatively simple failure pattern under uniaxial tension. As the laminate is deformed, all laminae which have been prestrained due to interlaminar residual fabrication stresses are subjected to the same external (in-plane) strain. In the initial stage of loading, the whole laminate deforms linearly. With continued loading, the strain (sum of prestrain and external strain) reaches a critical level and a translaminar failure by cracking parallel to fibers occurs in the 90° lamina. The failure of 90° laminae reduces their share of load and transfers it to the unbroken portion of the laminate. As a consequence, the stress-strain curve showed a marked reduction of laminate modulus above the yield point (Figure 1a).

During the testing it was observed that the successive occurrence of translaminar cracks (accompanied by acoustic emission) started in the 90° laminae slightly past the yield point. Delamination was rarely observed at the junction between the interlaminar region and the translaminar crack. When the crack density became relatively high in the 90° laminae, cracks started to form also in the 0° lamina. After "whitening" of all laminae by densely populated cracks, the laminate continued to deform until catastrophic failure occurred across

the fibers in the 0° lamina. At catastrophic failure, extensive de-
lamination took place between the 0° and the 90° laminae. The whole
failure pattern described above was not affected by water absorption
despite the change in the level of yield stress and ultimate strength.

Based on the observed behavior, it can be assumed that the
yielding of the 90/0/90° laminate is governed primarily by the failure
of 90° lamina. Since the 90° lamina is prestrained by the interlam-
inar residual stress and strained externally by the applied load with
the restraint of the 0° lamina, the following properties play a domi-
nant role in the determination of yield strain of the laminate:

(1) failure strain of unrestrained 90° lamina, i.e., trans-
verse ultimate strain of unidirectional lamina,

(2) degree of bonding between 90° and 0° laminae (which con-
trols the degree of load transfer between the laminae), and

(3) prestrain in the 90° lamina due to the interlaminar
residual stress.

If the interlaminar bonding is strong enough to preclude de-
lamination before the failure of 90° lamina, the yield strain can be
approximated as a difference between (1) the failure strain of 90°
lamina without restraint and (2) the prestrain in the 90° lamina,
assuming negligible Poisson's ratio effects (Figure 1b).

Therefore, the possible change in each of the above-mentioned
properties must be investigated to explain the change of yield strain
(or stress) by water absorption. So far the effect of water absorp-
tion on the interlaminar residual stress and the failure stress of 90°
lamina has been examined by separate experiments. The results are
presented in the following three sections.

A. Effect of Water Absorption on the Interlaminar Residual Stress
   of 90/0/90° Laminate

When a crossply laminate made up of 0° and 90° laminae is
cooled down from the cure temperature to room temperature, 90° laminae
having higher thermal expansion coefficients try to shrink more than 0°
laminae along the 0° direction. If the same dimensional contraction is
dictated by strong bonding between the laminae, each lamina will be in
a state of interlaminar residual stress. Although the isothermal con-
traction of each lamina due to the resin shrinkage in the curing
process can influence the residual stress, its effect is probably
negligible because of stress relaxation at the high curing temperature

(e.g., 163 C in E glass/1009 resin system); the interlaminar residual stress results mainly from the restraint against the differential contractive strain of 0° and 90° laminae during the cooling.

Interlaminar residual stress can be partially relieved by the distortion (warpage) of the laminate (9,10). In actual crossply laminates, the distortion is avoided by mid-plane symmetry (e.g., 90/0/90/0/90/0/90°). In this case the distortional forces exerted by all individual laminae will be balanced and no out-of-plane warping will occur. Therefore, in-plane shear stresses cannot occur in the individual laminae (11), and only normal stresses are possible: 90° lamina in tension and 0° lamina in compression along the 0° (load) direction. No warpage was in fact observed in the 90/0/90° symmetric laminates used in this study.

In this work, the differential contractive strain of 0° and 90° laminae was estimated two-dimensionally by measuring the degree of warpage in an unbalanced laminate. In order to obtain unbalanced 0/90° laminate specimens, the top laminae were removed from the 90/0/90° laminate specimens which had been immersed in water for various periods of time. The curvature of each 0/90° specimen was measured assuming the arc of the now warped specimen to be circular. The results showed that the curvature of 0/90° specimens decreased in a nearly linear fashion with increasing amount of water absorption (Figure 6).
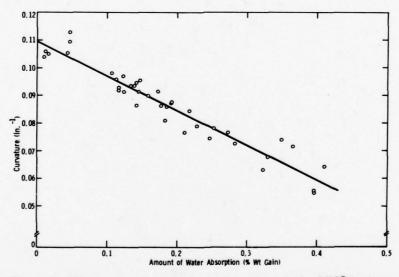

Figure 6. Effect of water absorption on the curvature of a 0/90° laminate.

There have been several equations derived to relate the curvature of unbalanced laminates to the differential thermal strain of the constituent laminae (9,10,12,13). The simplest of this is Timoshenko's bimetallic thermostat equation which was derived from the strain energy consideration (12). Although this equation ignores the local shear interaction between the laminae, it is shown to agree reasonably well with the more rigorous solutions (13) for a very thin laminate (with the thickness-to-length ratio less than 0.05). Since the 0/90° specimens of this study have a thickness-to-length ratio less than 0.005, Timoshenko's equation was used in the calculation of differential strain of 0° and 90° laminae (14). The following assumptions were made:

(1) linear elastic behavior of 0° and 90° laminae,

(2) perfect bonding between the laminae, and

(3) negligible Poisson's effect.

The calculation of differential strain requires values of tensile moduli of 90° and 0° laminae. The modulus of each 90° and 0° lamina which was removed from the immersed laminate specimens was measured and found to be relatively unchanged by water absorption up to the point of 0.4% weight gain (Figure 7). Using the average values of moduli of 90° and 0° laminae, the differential strain ($\Delta\ell/\ell_o$) of 0° and 90° laminae was calculated from the curvature of 0/90° specimens. Figure 8 shows that the differential contraction of 0° and 90° laminae due to cooling decreased linearly with increasing amount of water absorption. This phenomenon could be explained by one or more of three factors.
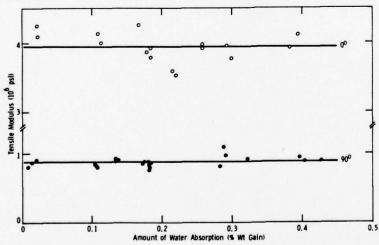

Figure 7. Effect of water absorption on the tensile modulus of 0° and 90° laminae.
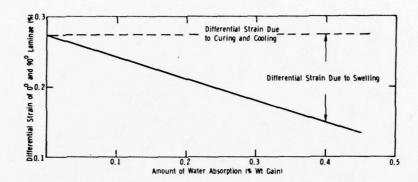
Figure 8. Effect of water absorption on the calculated differential strain of 0° and 90° laminae.

(1) Reduction in the modulus of 90° lamina due to the "plasticization" of matrix resin by water absorption. (Water molecules can disrupt intermolecular bonding forces of resin thus shifting its transition temperatures.)

(2) Reduction in the degree of interlaminar bonding.

(3) Development of differential expansive strain of 0° and 90° laminae by swelling of matrix resin.

The first factor can be dismissed based on the data which showed no significant change in the modulus of 90° lamina by water absorption (Figure 7). Presumably, the reduction in the glass transition temperature of matrix resin by water absorption has a negligible effect on the value of room temperature modulus up to the point of 0.4% weight gain (1.6% weight gain in the resin phase assuming no water absorption of other components).

In the unbalanced laminate, general or intermittent loss of interlaminar bonding can partially relieve interlaminar residual stress, thereby resulting in a lower degree of warpage. Loss of interlaminar bonding by water absorption is theoretically possible through debonding between the fiber (of 0° lamina closest to the interlaminar region) and resin at the interlaminar region. However, with weaker interlaminar bonding, it is also possible for the cracks to propagate more easily along the interlaminar region under external force. In the laminate system reported here, no increasing tendency toward coupling between the delamination and the translaminar crack formation could be observed with increasing amounts of water absorption. This fact discredits indirectly a possible contribution of the second factor to the reduction of calculated differential strain of the 0° and 90° laminae.

With present information, the third factor is most likely to explain the reduction of differential contraction. Ishai observed anisotropic volume expansion behavior of unidirectional composites by water absorption (15). According to his results, a glass fiber/epoxy resin lamina shows a considerable volume expansion perpendicular to the fiber direction by water absorption. Likewise, the same lamina shows a negligible amount of volume expansion along the fiber direction. Based on his data, it can be postulated that the differential expansion of 0° and 90° laminae due to swelling has the opposite effect on the differential contraction due to cooling.

Without taking into account the differential expansion due to swelling observed by Ishai, the reduction of interlaminar residual stress arising from the reduction of differential contraction of 0° and 90° laminae by water absorption can be estimated. The interlaminar residual stress was calculated two-dimensionally using a simplified model by Thompson (14,16). His analysis showed that the interlaminar residual stress in the 90° lamina along the 0° direction (as an integrated average value through the lamina thickness) is given by:

$$\sigma_{r\ 90°} = [(\Delta \ell/\ell_o)\ E_{90°}]/1 + (E_{90°}\ A_{90°}/E_{0°}\ A_{0°})$$

($\Delta \ell/\ell_o$ = differential strain of 0° and 90° laminae, E = modulus, A = cross-sectional area). The tensile prestrain ($\sigma_{r\ 90°}/E_{90°}$) in the 90° lamina could be calculated from the interlaminar residual stress ($\sigma_{r\ 90°}$) in a straightforward manner. The tensile prestrain in the 90° lamina calculated as above showed a linear decrease with increasing amount of water absorption. It reflects clearly the relaxation of interlaminar residual stress by water absorption.

B. Effect of Water Absorption on the Failure Stress of Unrestrained 90° Lamina

Under uniaxial tension, the 90° lamina specimen deforms linearly in the initial stage; however, it later exhibits nonlinear deformation before final failure (Figure 1a). The results based on testing of 90° lamina removed from the laminate show that the magnitude of nonlinearity increased substantially by water absorption (Figure 9). On the other hand, the failure stress (or strain) of 90° lamina showed a broad maximum around the point of 0.2% weight gain, as in the case of yield stress of 90/0/90° laminate. The complete explanation of these phenomena will comprise future work ("The Effect of Static Immersion in Water on the Tensile Strength of a Lamina").
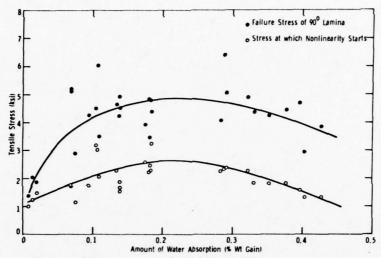
Figure 9. Effect of water absorption on the failure stress of a
90° lamina without the restraint of a 0° lamina.

C. Prediction of the Change in Yield Stress of 90/0/90°
    Laminate Induced by Water Absorption

        In the two previous sections, the effect of water absorption
on the interlaminar residual stress of 90/0/90° laminates and the
failure stress of unrestrained 90° lamina was discussed. Based on a
two-dimensional model, the tensile prestrain in the 90° lamina due to
the interlaminar residual stress was predicted to decrease continuously
with increasing amount of water absorption. The change in tensile
failure stress or strain of 90° lamina without the restraint of 0°
lamina by water absorption was also observed. The next task is to
predict the change in yield stress of a 90/0/90° laminate based on
the above information.

        As shown in the earlier discussion, if the interlaminar bond-
ing is perfectly strong (no delamination before lamina failure), the
yield strain of 90/0/90° laminate can be approximated as $\varepsilon_f\ 90° - \varepsilon_r\ 90°$,
assuming negligible Poisson's effect (Figure 1b). The yield stress of
90/0/90° laminate can be obtained from the values of $\varepsilon_f\ 90°$, $\varepsilon_r\ 90°$,
and $E_{90/0/90°}$ (Figure 10). Although the values of $\varepsilon_f\ 90°$ could be
estimated directly from the stress-strain curve, the calculated values
of $\sigma_f\ 90°/E_{90°}$ were used instead to minimize the effect of nonlinearity.

        The calculated change in yield stress of 90/0/90° laminate by
water absorption is shown as a dotted line in Figure 4. While based on
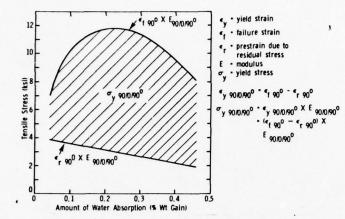limited data, the prediction of the change in yield stress is relatively

Figure 10. Effect of water absorption on the calculated yield stress of a 90/0/90° laminate.

close at lower amounts of water absorption. But with greater water absorption, the experimental data points deviate from the prediction somewhat. Aside from the experimental errors (weight gain, curvature of 0/90° specimens, failure stress of 90° lamina, etc.), two factors are considered to be responsible for the deviation:

(1) higher degree of nonlinearity in the stress-strain curve of 90° lamina with increasing amounts of water absorption, and

(2) possible change in the Poisson's ratio of 90° lamina by water absorption.

CONCLUDING REMARKS

This report describes studies to determine the effect of water absorption on the tensile strength of 90/0/90° laminates. The ultimate strength was found to decrease continuously with increasing amounts of water absorption. Yield stress first increased and then decreased to form a broad maximum. Investigation focused on the question of how and why the yield stress shows a maximum at a certain amount of water absorption.

Based on the curvature measurements of unbalanced laminates and the two-dimensional theoretical model, a prediction was made of the relaxation of interlaminar residual stress in the laminate by water absorption. The change in failure stress of an unrestrained 90° lamina arising from water absorption was also observed experimentally. Assuming perfect bonding between the laminae, the change of yield stress could be predicted by overlapping the decrease of interlaminar residual stress (as a prestress) and the change of failure stress of 90° lamina.

It was found that the relaxation of interlaminar residual stress by water absorption does not result from the reduction of modulus of 90° laminae. Presumably, the reduction in the glass transition temperature of the matrix resin by water absorption (plasticization) has negligible effects on the value of room temperature modulus. Anisotropic volume expansion of the laminae by water absorption is considered to be responsible for the relaxation of interlaminar residual stress.

REFERENCES

1. SACHER, R. E., and ABRAMO, E. J. Proc. 32nd Ann. SPE Tech. Conf., 1974, p. 415.

2. BLAGA, A. Polymer Eng. Sci., v. 12, no. 1, 1972, p. 53.

3. KAMAL, M. R. Polymer Eng. Sci., v. 10, no. 2, 1970, p. 108.

4. HAHN, H. T., and TSAI, S. W. J. Composite Materials, v. 7, no. 102, 1973.

5. HALPIN, J. C. *Glass Reinforced Epoxy Systems (Materials Technology Series Vol. 2)*. C. J. Hilado, ed., Technomic Publishing Co., Westport, CN, 1974.

6. American Society for Testing and Materials Standard D792-66.

7. LEE, H., GORALESKI, E., and EAGLE, C. V. Proc. 2nd National SAMPE Tech. Conf., 1970, p. 503.

8. ROTEM, A., and HASHIN, Z. J. Composites Materials, v. 9, 1975, p. 191.

9. CHAMIS, C. C. Proc. 30th Ann. SPI Reinf. Plastics/Composites Inst. Tech. Conf., sect. 18-C, 1975.

10. EPSTEIN, M. M., COOPER, C. W., STICKNEY, P. B., and BELL, J. C. Appl. Polymer Symposia, v. 4, 1967, p. 219.

11. SCHNEIDER, W. Kunststoffe, v. 61, 1971, p. 273.

12. BRAND, R. H., and BACKER, S. Textile Res. J., v. 32, 1962, p. 39.

13. CHOW, T. S. J. Appl. Physics, v. 47, no. 4, 1976, p. 1351.

14. LEE, B. L., LEWIS, R. W., and SACHER, R. E. Army Materials and Mechanics Research Center, Technical Report in process.

15. ISHAI, O. Polymer Eng. Sci., v. 15, no. 7, 1975, p. 486.

16. THOMPSON, B. UARL Report J213186-9.

A METHOD FOR THE RAPID EVALUATION OF HAZARDS
FROM TOXIC WASTE DEPOSITS (U)

DUANE E. LONG, MR.
*MARTIN J. HOULE, MR.
CHEMICAL LABORATORY DIVISION, DUGWAY PROVING GROUND
DUGWAY, UTAH 84022

Growing public and governmental concern for the overall dete-
rioration of the environment has led to laws regulating the release of
potentially toxic materials into the air or water. However, the result-
ing pollution control processes also produce large quantities of
potentially hazardous wastes.   The Department of Defense, a leader in
pollution control and site reclamation, is concerned about monitoring
for the possible release of toxic components from wastes produced at
installations engaged in explosives manufacturing, electroplating
operations, chemical weapons demilitarization, etc. Toxic materials
may be present in the liquid fraction of a waste, or in the leachate
produced by rain or surface water percolating through a solid waste.
As these solutions soak through the soil they eventually can reach the
ground water and present a hazard to water users.

The migration of chemical substances through soil is usually
determined in the laboratory using columns packed with soil to a pre-
determined bulk density. These soil columns are challenged with a
solution extracted from a waste by water or some other solvent such as
municipal landfill leachate, or the soil is treated with simple solu-
tions of the ion under study. A useful configuration is shown in
Figure 1, along with typical plots of the data obtained from continu-
ously-leached columns. (This example illustrates the case where the
concentration of the compound of interest is reduced by passing through
the soil.) Besides requiring close attention to regulating flow-rate,
an important limitation of continuously-leached column experiments is
the time and effort required to obtain and analyze a sufficient number
of samples to make predictions of migration rates and of the toxic haz-
ards that could result from leaching the waste. This usually requires
months and may even take years, depending upon the flow-rate of the
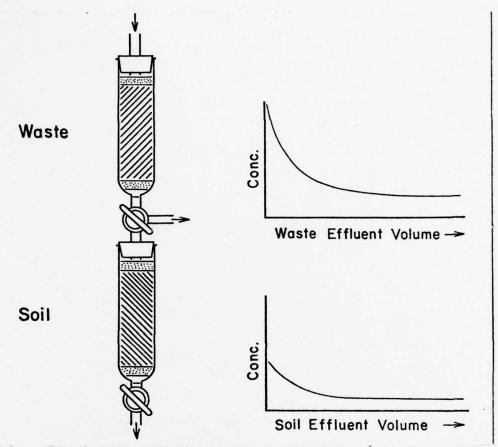leaching solution through the column. The information obtained from a

Figure 1. Continuously-leached columns and associated output plots.

relatively short-term column study cannot be expected to describe what will occur during years of leaching.

An Army activity wishing to evaluate the potential hazard from the disposal of a waste is faced not only with the problem of obtaining valid results in a reasonable length of time but also with designing experiments that adequately represent the field situation. An investigator must select values for each experimental parameter such as leaching solvent flow-rate, head pressure, soil bulk density, column diameter, waste-to-soil ratio, leaching time, etc. The choice of these values may not all be entirely arbitrary, but a given set will yield results which probably apply only to that particular combination of conditions and the experiment may not be very useful for making general predictions of the migration of chemical compounds through soil.

By having a more rapid and flexible experimental approach, a wider range of conditions can be investigated within the framework of

factorial experiment designs which do allow making predictions even in the presence of statistical interaction between multiple variables. A fast method also allows making timely determinations, on demand, for each specific situation.

A graded serial batch procedure which is rapid and widely applicable has been developed in this laboratory. To validate this procedure, it first was necessary to establish a correlation between the batch procedure and continuously-leached columns. Experimental comparisons obtained were good and a consideration of the plotting parameters (discussed below) showed that this new approach could accelerate the testing of waste leachability and contaminant movement through soils.[1] The procedure also normalizes the results so they can be correlated to a range of field conditions. Some elements of this study have been corroborated by investigations conducted elsewhere.[2,3]

CORRELATING CONTINUOUS AND BATCHWISE LEACHING

The data obtained from continuously leached columns may be presented in several ways. One technique is to plot the concentration of the chemical of interest found in the waste or soil column sample versus the cumulative volume through the column. The common way of expressing the cumulative volume is to use the cumulative pore volume calculated for the type and weight of soil employed. (The pore volume is the interstitial void in a volume of soil, and the total void space depends upon soil type and mass.) The scale of the cumulative volume axis therefore changes for different soil types and sample sizes when pore volume is employed. Figure 2 is an example showing the difference obtained with pore volumes of 40 and 60 milliliters. The corresponding total volume in milliliters is appended for comparison.

It often is not practical or possible to determine a pore volume for a waste due to its physical form (heterogeneous suspension, liquid, etc). Using the soil column pore volume as the measure of liquid volume through the waste allows correlating the waste-column output with the soil-column results in a given set of experiments. However, instead of using the soil pore volume as the principle plotting parameter, it is much more flexible to plot the observed concentration of a chemical in an extract versus the cumulative milliliters of leaching solvent per gram of waste or soil. This makes the scaling independent of soil type, soil sample weight, and waste-to-soil ratio, and allows the direct comparison of many different designs of experiments. The area under the curve represents the total weight of a chemical extracted per gram of waste or soil.

Batchwise extractions can be related to continuously-leached columns by recognizing that continuous leaching is equivalent to running a series of discrete extractions spaced by the frequency of collecting the effluent sample. Figure 3 shows that the concentration of the periodic column samples can be plotted to represent the average for
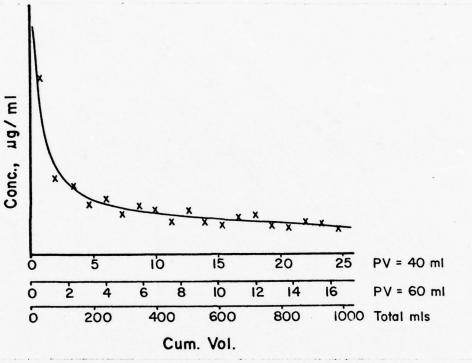
Figure 2. Differences in scales used to plot cumulative volume.

that sampling period. Thus, samples from the continuous leaching of a column correspond to sequential batchwise extractions by volumes of extractant equal to the volume passing through a column between the taking of samples.

When extracting a batch of waste or soil, instead of using the same volume of solvent for each successive extraction, the solvent-to-waste or -soil ratios can be graded in size as indicated by the extraction volumes pictured in Figure 4. A small solvent-to-solids ratio should probably always be employed for the first extractions; this is when the soluble species will be the most highly concentrated in the extract and the ionic strength will be at its maximum. Greater dilutions would reduce this, possibly affecting the solubility of other components. After the more soluble components have been extracted, the solvent-to-solids ratio can be greatly increased, thus reducing the total number of extractions required. The further along the cumulative milliliters per gram axis that the extraction volumes extend, the longer the period of column leaching the batch work is equivalent to.

Since batch extractions are quite rapid compared to letting the liquid percolate through a column, sequential batch extractions can be the basis for accelerated testing of wastes and soils. By relating
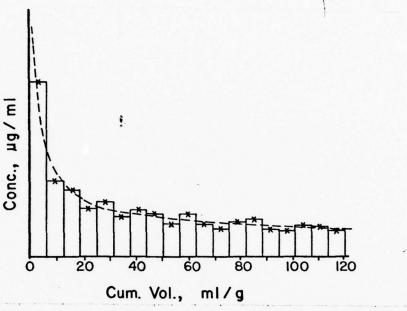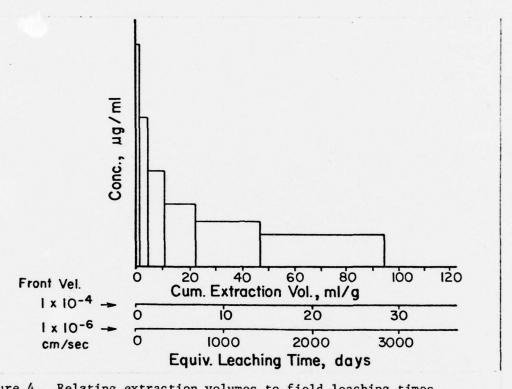
Figure 3. Relation between batch extractions and continuously-leached columns.

the rate of liquid-front movement to the volume flow rate per unit time, a scale of equivalent contact time can be added parallel to the axis labeled cumulative volume per gram, as shown in Figure 4. (The development of this concept is discussed in detail elsewhere.)[1] Table 1 lists the cumulative volume, in milliliters per gram, for the liquid-to-solid ratios employed, together with the equivalent exposure time for liquid front velocities of 1×10^{-4}, 10^{-5}, and 10^{-6} centimeters per second.

THE GRADED SERIAL BATCH EXTRACTION OF WASTES AND SOILS

The waste composition changes as components are leached from the waste. Each succeeding portion of extract will therefore generally have a different composition. Besides being challenged by a changing solution, the soil's ion-removal characteristics continually change with time as the soil becomes conditioned and loaded by the passage of waste extracts. Since each portion of waste is changed by passage through a segment of soil, the conditioning each succeeding segment of soil receives is different and each segment therefore may remove different proportions of the various ions present in the waste extract. So although the soil segments start out the same, in effect they become different soils due to the passage of the different waste extracts.

The soil removes ions from the waste, but the waste extract can also displace ions from the soil. In addition, soil can pick up a

Figure 4. Relating extraction volumes to field leaching times.

TABLE 1.

Correlation Between Extraction Volume and Penetration Time

Extraction Number	Water Added, ml/g	Cumul ml/g	Equivalent Days of Penetration[a]		
			10^{-4}	10^{-5}	10^{-6} cm/sec
1	2	2	0.6	6	60
2	3	5	1.5	15	150
3	6	11	3.3	33	330
4	12	23	6.9	69	690 (1.9 yr)
5	24	47	14.1	141	1410 (3.9 yr)
6	48	95	28.5	285 (.78 yr)	2850 (7.8 yr)
7	96	191	57.3	573 (1.6 yr)	5730 (15.7 yr)

[a]At the specified liquid front velocity through a typical soil having bulk density of 1.6 g/cc and a pore volume of 0.24 ml/g.

specific ion from a waste solution of one composition and then give it up again as the liquid composition changes. The soil may also give up ions later because of intervening conditioning of the soil by the passage of the changing waste extract solution.

If extract samples were taken within a layer of soil, it would be possible to study this dynamically-changing situation. This can be accomplished by placing sampling ports in the side of a soil column, as shown in Figure 5. The same result can be attained in a shorter time with far fewer equipment difficulties by putting waste extracts on successive batches of soil and taking a sample after each extraction. A batch of soil then will represent a segment of soil from a soil layer.

Normally, the distribution of substances retained by the soil column is determined after leaching is conducted and the soil column is
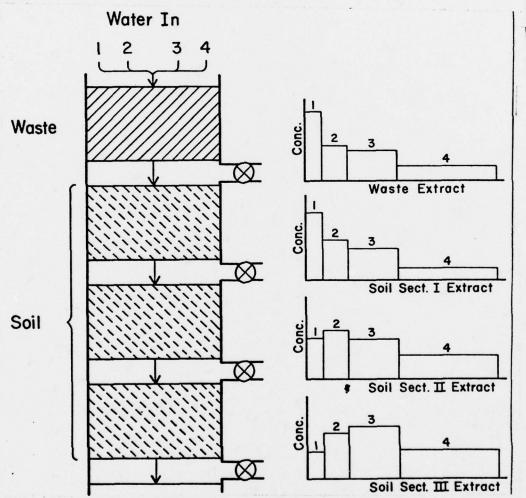

Figure 5. Challenging multiple soil segments with successive extracts of waste.

sectioned and analyzed. But, a serial batch approach, with sampling between batches of soil, yields a real-time picture of what is happening within a bed of soil and provides data which could allow extrapolating to the effect of thicker strata--something which cannot be done with validity from experiments with only a single layer, or from experiments which use artificially simple solutions. It is re-emphasized that batchwise testing also yields its information in a small fraction of the time required by columns or field studies.

THE EXPERIMENTAL PROCEDURE

A sequence of seven extracts was made from each type of industrial waste. First, a sample of waste was dried to determine moisture content, then sufficient undried sample to give 300 grams dry weight was weighed into a 2-quart, wide-mouth screw cap jar. (Drying the sample could affect hydrated species and drastically reduce the solubility. If the waste had supernatant water, the volume of the water was considered as part or all of the first extraction.) Appropriate volumes of water were added for each extraction to produce the liquid-to-solid ratio given in the second column of Table 2. The bottle was shaken gently four or five times each day. (Continual mechanical shaking was not used because of concern that it might abrade the waste agglomerates, making them more susceptible to extraction.) The time required to reach equilibrium can be determined by periodically withdrawing an aliquot for analysis; 24 hours is adequate for most wastes of small particle size. At the end of each extraction period, the mixture was filtered under vacuum using a hardened filter paper (such as Whatman 54) in a Buchner funnel. An aliquot of approximately 20

TABLE 2.
Specifications for Serial Batch Extractions

Extraction Number	Water Added, ml/g	Volume of Water, ml, Extracting 300 g Waste	Volume of Filtrate Onto a Soil		
			I 60 g Soil	II 30 g Soil	III 15 g Soil
1	2	600	120	60	30
2	3	900	180	90	45
3	6	1,800	360	180	90
4	12	3,600	720	360	180
5	24	7,200	1,400	720	360
6	48	14,400	2,800	1,400	720
7	96	28,800	5,760	2,280	1,440

milliliters was withdrawn for analysis and filtered through a 0.5 μ Millipore filter to remove fine particulates which might have bypassed the filter paper to possibly dissolve when the sample was acidified.

(After measuring conductance and pH, one percent concentrated nitric acid was added to inhibit precipitation while standing.) The solid waste residue was transfered back to the jar and mixed with the volume of water specified for the next extraction.

In the procedure detailed here, the liquid-to-solid ratio was continually increased to further accelerate the testing--the volume of each extraction after the second one was made double the one before, which redoubles the time represented by that extract. With some wastes adequate results may be obtainable from using very large volumes right from the first (or one or two extractions using small liquid-to-solid ratios, followed by a very large one) but this would have to be checked for each kind of waste. However, this procedure will allow rapid simulation of long leaching periods and could be useful in the routine monitoring of variations in waste composition and leachability.[4]

The filtrate resulting from each sequential extraction of the waste was mixed with the first of three batches of each kind of soil. The weights of soil used were 60, 30, and 15 grams, representing sections I, II, and III, respectively. This gradation in weight allows taking an aliquot of the extract for analysis and having enough left over to challenge the next soil batch with the same liquid-to-solid ratio. Extracting 300 grams of waste yields sufficient solution to challenge three different kinds of soil in experiments set up with the proportions stated in Table 2.

Although the soil equilibrates in 6 hours or less[5], each solution was kept in contact with the batch of soil before filtration for the same length of time as used to extract the waste. This was to keep the samples progressing smoothly without gaps in the series. After filtering the soil extract, an aliquot was refiltered through Millipore and saved for analysis. The appropriate volume of the remaining filtrate was added to the next batch of soil. The soil exposed to the first waste extract was recovered and mixed with the second waste extract in the series. This was repeated until the waste had been extracted seven times and each waste extract had progressed through all three soil batches. This procedure was run in duplicate.

APPLICATION OF THE BATCH TECHNIQUE TO INDUSTRIAL WASTES

Two industrial wastes of widely divergent characteristics were carried through a series of seven batch extractions. The resulting extracts were used to challenge three different kinds of soils. The extract samples were analyzed to determine the amount of each ion of interest, the pH, and the conductivity at every stage of the batch tests. In about two weeks of laboratory work with each waste, these experiments simulated approximately 1.6 years of leaching under field conditions that give a liquid front velocity of 1×10^{-5} cm/sec. An eighth extraction would have extended this to the equivalent of three years in the field. The composition of the wastes and soils are

detailed below, together with the interpretation of the experimental results.

WASTE COMPOSITION

      a. <u>Elemental Phosphorus Production Waste</u>. This waste originated from the production of elemental phosphorus by the electric furnace method. First, the ore is dried and calcined in a kiln and gases containing phosphorus, fluorides, and fuel decomposition products are emitted. These gases are passed through a water scrubber and the resulting liquor is treated with lime to precipitate these compounds. The solid fraction of the waste is composed primarily of calcium phosphate, calcium fluoride, calcium sulfate, and unreacted lime. The species of interest in the waste leaching and soil migration study were inorganic phosphorus (probably present as phosphates) and fluoride. These anions were present in a slurry of high alkalinity (pH 12.7).

      b. <u>Zinc-Carbon Battery Rejects</u>. This waste consists of broken-open reject batteries. (Approximately one percent of the batteries produced are rejected.) The extract samples were analyzed for the mercury, zinc, cadmium, and lead leached from the batteries. These cations were present in a solution only slightly above neutrality (pH 7.5).

      c. <u>Soil Composition</u>. Three soils were investigated for their ability to remove the metals of interest from extracts of the two wastes. The soils chosen were Chalmers (a gray, silty, clay loam from Indiana, a Mollisol), Davidson (a red clay from North Carolina, an Ultisol), and Nicholson (a yellow silty clay from Kentucky, an Alfisol) These soils were selected because of differences in their chemical properties and clay mineralogy. Chalmers and Nicholson soils have similar surface areas but the higher percentage of clay in the Nicholson soils yields a higher cation exchange capacity. In addition, the clay mineral composition is much different. The Chalmers clay composition is largely montmorillonite, with very small amounts of vermiculite, chlorite, and kaolinite. The Nicholson clay fraction is predominately vermiculite with only a trace of mica and kaolinite. In contrast, Davidson soil has a low surface area and cation exchange capacity. The clay fraction is predominately kaolinite which has a significantly lower cation exchange capacity than the above clays but this soil contains a higher percentage of hydrous oxides of iron. It has been shown that iron oxides play a major role in heavy and trace metal removal.[6]

      d. <u>Interpretation of Results</u>. Plotting the waste output as in Figure 5 shows the depletion of the waste with continued leaching. To show the effect of passing the waste extract through soil, the histogram of Figure 6 is more useful. It presents the results obtained from extracting a batch of waste and placing these solutions in succession on three batches of soil. (Multiplying the observed concentration in micrograms/milliliter by the individual batch extraction volume
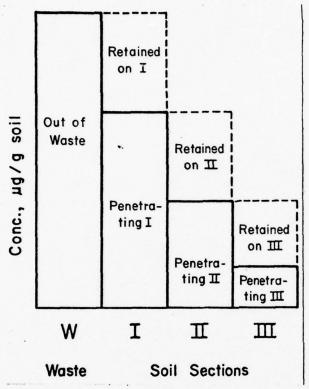
Figure 6. Histogram showing the penetration and retention of a species by soil.

per gram, milliliter/gram, converts the values to micrograms/gram of waste or soil.) The height of the histogram bar labeled W represents the micrograms of, e.g., phosphorus extracted per gram of waste. This is the challenge to the batch representing soil section I. The height of the bar labeled I shows the concentration of phosphorus penetrating the batch representing soil section I, and the difference in height between I and W is the amount of phosphorus retained per gram of soil. The ratio of (W-I) to W is the fraction removed by that soil section. Similarly, I is the challenge to II, and bar II shows the penetration through II. If the fraction of phosphorus removed by each soil section is different, this shows that the removal characteristics of the soil are affected by conditioning and by changes in the extract. This can be further studied by comparing the histograms for the different extractions.

The soil batches can be treated in pairs as above, or it can be considered that a given amount of waste has challenged three different amounts of soil: section I, section I + II, and sections I + II + III. Thus, the fraction removed can be calculated for three different waste-to-soil ratios.

Because the fraction removed corresponds to the ratio of the concentration in the soil to the concentration in the solution, this is equivalent to the distribution coefficient, which is the slope of the adsorption isotherm. By calculating the fraction removed for each histogram bar, it is possible to follow the change in distribution ratio (the change in slope of the adsorption isotherm) for each different waste-to-soil ratio and for each serial extraction (which show the effect of the changing sample matrix). This is of considerable

importance for modeling and for making predictions of contaminant move-
ment through soils.

Figure 7 is a composite plot of histograms which show the
extraction of inorganic phosphorus from phosphorus production waste and
its penetration through Davidson and Nicholson soils. (Chalmers gave
results similar to Davidson so it is not plotted here.) Although flu-
orine is also of interest in this waste, phosphorus is used here as the
example. Because of space limitations, only four of the seven serial
extractions are pictured. Going vertically down the figure viewing any
single column shows how the extraction and penetration of phosphorus
changes as the leaching progresses. For example, the W's show a
continuous decrease in the phosphorus concentration in the waste
extract as the leaching continues.

It is seen that neither soil is effective in removing inor-
ganic phosphorus from this waste extract. (When the histogram bar is
higher than the preceding one, it shows that the soil is releasing
phosphorus.) Davidson is slightly more effective in retarding the
phosphorus movement than is Nicholson soil, but the phosphorus appears
to move through the soil as a zone similar to that obtained in elution
chromatography. This is probably due to the high pH of the waste
extract (the pH ranged from 12.7 to 11.6 in the four batches). This
drastically changed the environment within most of the soil batches by
changing the soil pH from slightly acid to strongly basic. This condi-
tion favors the mobility of anionic species.

The soils had a much different effect on the waste extract
from the zinc-carbon batteries than on the phosphorus waste extract.
Although several metal ions were found in significant concentrations in
the battery waste leachate, zinc will be used here as the example. Fig-
ure 8 presents a composite plot of histograms giving the penetration of
zinc in battery extract through Davidson and Nicholson soils. (Chalmers
soil gave nearly the same results for zinc as did Davidson soil.) Both
soils removed significant quantities of zinc from the waste extract,
but Davidson soil is considerably more effective than is Nicholson soil.
Although Nicholson soil has much higher surface area and cation
exchange capacity than Davidson soil, the iron oxide content of the
Davidson soil is much higher than the Nicholson. This is one of the
important factors in removing zinc. The pH of the soil extracts was
also affected. Even though the first waste extract had a pH of 7.6,
the Nicholson soil extracts from batches I,II, and III were quite acid:
pH 5.8, 4.5, and 4.7, respectively, for the first extract. In compari-
son, the first Davidson soil extracts were pH 6.4, 6.3, and 6.4. The
pH of the extracts from both soils were not comparable until they had
been exposed to the fourth waste extract. The ability of Nicholson soil
to remove zinc began to approach the removal capability of Davidson
soil by the fourth in the series of extractions, either because of a
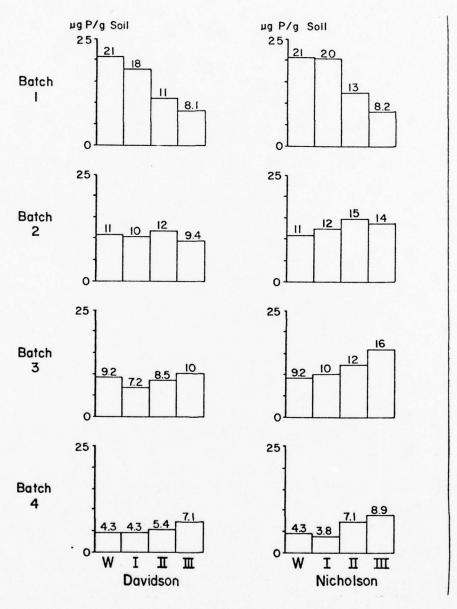conditioning of the Nicholson soil or because the zinc is then present

Figure 7. The extraction and penetration through soils of inorganic phosphorus from phosphorus production waste.
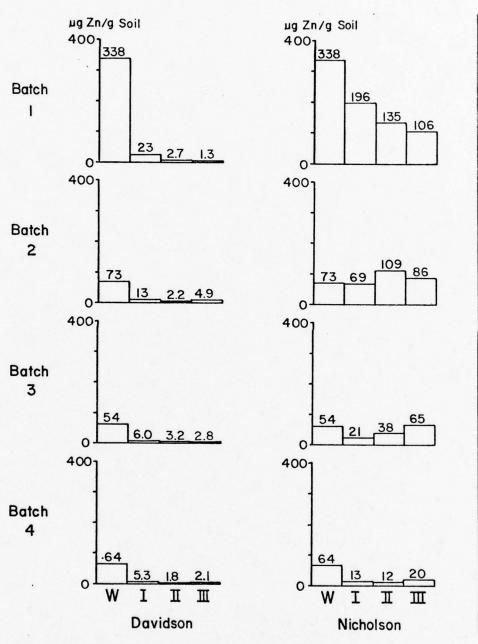
Figure 8. The extraction and penetration through soils of zinc from zinc-carbon batteries.

LONG & *HOULE

in a waste leachate of different composition. Release of zinc from a soil section is shown whenever one histogram is higher than a preceding bar. Nicholson soil shows a considerable release of zinc from the second and third soil batches. This is probably zinc taken out of previous extracts but later released by the soil due to the change in composition of the later extracts and the different soil history.

CONCLUSION

It has been demonstrated that the leaching of a waste can be characterized and the ability of a soil to remove a chemical species from the waste leachate can be rapidly evaluated using serial batchwise extractions. This new method for evaluating a waste-soil system is not only much faster, it gives even more information than columns about the pickup and release characteristics of soils for toxic species in a dynamically changing situation. The batch technique is also far more convenient for rapidly investigating the effect of various environmental factors and lends itself to evaluating the effects of other variables such as sunlight, drying or freezing cycles, etc., via factorial experiments.

No one method can give all the answers, but the graded serial batch approach can quickly provide much of the information needed to assess the categories of hazards from a class of wastes and to make decisions concerning the suitability of soil types for inhibiting the migration of hazardous chemicals.

REFERENCES

1. Houle, M. J. and Long, D. E., In: Land Disposal of Hazardous Waste, Proceding of the Fourth Annual Research Symposium, Mar 78, US Environmental Protection Agency, Cincinnati, OH 45268, In Press.
2. Liskowitz, J. W., et al., In: Residual Management by Land Disposal, Proceedings of the Hazardous Waste Research Symposium, Feb 76, EPA-600/9-76-015, US Environmental Protection Agency, Cincinnati, OH 45268, p. 162-176.
3. Farquar, G. L., and Rovers, F. A., In: Gas and Leachates from Landfills: Formation, Collection, and Treatment, Proceedings of a Symposium at Rutgers University, New Brunswick, New Jersey, Mar 75, EPA-600/9-76-004, US Environmental Protection Agency, Cincinnati, OH 45268, p. 54-70.
4. Houle, M., Long, D., Bell, R., Weatherhead, D., and Soyland, J., In: Proceedings of a National Conference About Hazardous Waste Management, Feb 77, Prepared for US Environmental Protection Agency, In Press.
5. Griffin, R. A., and Shimp, N. F., Sep 75, US Environmental Protection Agency, Cincinnati, OH 45268, Contract No. 68-03-021. In Press.
6. Jenne, E. A., In: Trace Inorganics in Water, Advanced Chemistry Series 73: 337-387, 1968.

DYNAMICS OF PIN PALLET RUNAWAY ESCAPEMENT (U)

GERARD G. LOWEN, Professor
City College of New York, New York, NY 10031
*FREDERICK R. TEPPER, PhD
ARRADCOM, Dover, N.J. 07801

1. INTRODUCTION

This paper reports on the development of a simulation of a constant input torque pin pallet runaway escapement by the present authors [1][1]. The resulting program provides a basic tool for the analysis and synthesis of various safing and arming devices.

The dynamics of the regimes of motion of the escapement have been formulated. Coupled motion, with continuous contact between escape wheel tooth face and pallet pin, impact of the pin on the tooth face and uncoupled, or free, motion of these mechanism components were considered. The associated regime equations apply to entrance as well as exit conditions. Sensing expressions for the determination of the instantaneous positions of the pallet pin and the escape wheel form the basis of the controls of the computer program. In addition, the sensing equations indicate the presence of such pathological conditions as tip or back face contact. The simulation has been applied to the timing mechanism of the M525 fuze. The influence of changes in such parameters as escape wheel input torque, pallet moment of inertia, center distance, pallet radius, etc., on the mechanism delay time have been explored in detail by appropriate computer runs. The results, which are shown here, compare favorably with existing experimental data.

This effort represented an extension of the work of M. E. Anderson and S. L. Redmond [2]. New methods of contact kinematics for the coupled motion, of contact sensing, and of computational

[1] Numbers in brackets refer to refences in section 7.

controls were developed.

2. DESCRIPTION OF MOTION REGIMES AND SENSING TOOLS

The following describes the various dynamic regimes of the simulation, together with the applicable sensing parameters.

A. COUPLED MOTION

Figure 1 shows the entrance pallet pin as it is driven in coupled motion by tooth No. 1 of the escape wheel. The escape wheel angle ϕ is defined by the line from the escape wheel pivot O_s to the tip of the contacting tooth (or the one about to make contact) and the center line connecting O_s to the pallet pivot O_p. Similarly, the angle ψ, which is defined by the line from O_p to the active pallet pin center (entrance or exit) and the center line, describes the motion of the pallet. The escape wheel is driven by the constant moment T in the positive direction of rotation. While it is assumed that friction acts on the pallet pin-escape wheel tooth interface, it is neglected at both pivot pairs since investigation showed that its effects are negligible when the pivots are of the usual small diameter. The mathematical description of coupled motion is furnished by a non-linear second order differential equation in the escape wheel angle ϕ. The presence of the pallet is reflected back to the escape wheel in terms of a position dependent moment of inertia.
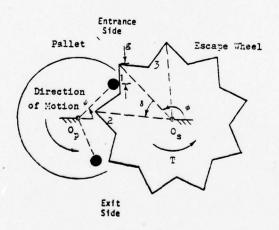
Figure 1 Coupled Motion

The quantity g, which represents the distance from the contact point to the tip of the escape wheel tooth, is used to determine the progress of coupled motion. Once the angle ϕ has been determined from the solution of the differential equation, both ψ and g may be computed from appropriate kinematic relationships.

B. FREE MOTION

When coupled motion is completed, i.e. g=0, or when separation of contact occurs after impact, escape wheel and pallet
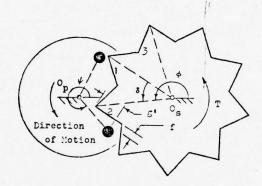

Figure 2 Free Motion

move independently of each other in free motion. Figure 2 shows this free motion for the exit phase of the action, i.e. the exit pallet pin is about to make contact with tooth No. 2 of the escape wheel. The constant torque T continues to act on the escape wheel, while the motion of the pallet depends only on its initial conditions. Again, any frictional retarding moments of the pivots are neglected. The motions of both components are represented by simple linear 2nd order differential equations. During this regime, position sensing is accomplished with the help of the quantities g' and f. The first of these quantities represents the distance of the pallet pin center from the tip of the escape wheel tooth, while the second measures the distance between the pallet pin and the tooth face. The solutions to the individual differential equations provide the angles ϕ and ψ. These are used to determine the above sensing parameters.

C. IMPACT

Impact follows free motion when both f=0 and g'<0,

and when the relative velocity between the contact surfaces warrants
it. (g' is negative because of the choice of the coordinate system.)
Such an impact usually reverses the motion of the pallet (see Figure
3) and under certain circumstances also temporarily reverses the
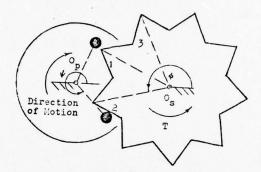motion of the escape wheel.
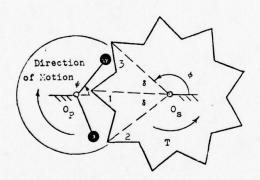
Figure 3 Exit Impact

Figure 4 Impending Entrance Impact

Figure 4 shows free motion for the subsequent entrance phase of the
mechanism, i.e. the top pallet pin is about to make contact with
tooth No. 3 of the escape wheel. The simulation recognizes only
contact on the front faces of the escape wheel teeth. This means
that such pathological conditions as impact on the tips or on the
back faces of the teeth are not considered. (The control quantities
g and g' make it clear when such a condition exists and the compu-
tation can then be discontinued.)

3. ESCAPEMENT NOMENCLATURE

Figure 5 shows a schematic representation of the pin pallet escapement and indicates its basic geometric nomenclature.
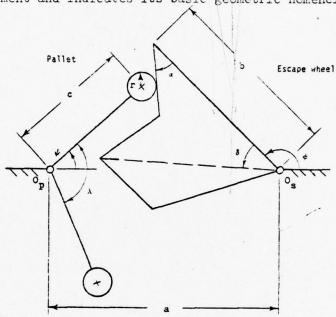

Figure 5 Escapement Nomenclature

a = Distance between pivots O_p and O_s
b = Escape wheel radius
c = Pallet radius (equal on entrance and on exit side)
r = Pallet pin radius (equal on entrance and on exit side)
α = Escape wheel tooth half angle
δ = Angle between escape wheel teeth
λ = Angle between pallet radii
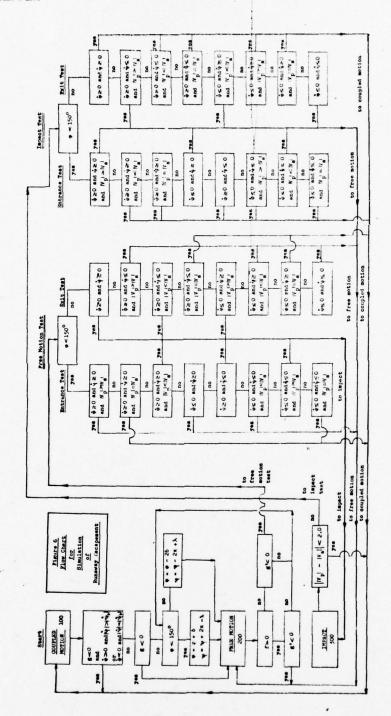
4. DESCRIPTION OF COMPUTER PROGRAM

The following gives the essential steps of the computer program[2]. Figure 6 represents the associated flow chart. The main program starts the simulation with coupled motion on the entrance side for a starting angle $\phi = 135°$. The total escape wheel angle is set

---

[2] The program shown is written in FORTRAN for the CDC System at ARRADCOM, Dover, NJ.

5

Figure 6
Flow Chart
for
Simulation
of
Runway Impingement

to 0°. This choice of starting angle represents the approximate midpoint of the coupled motion of the example mechanism. The computation is terminated either when t = .1 seconds, or when the total escape wheel angle equals 310°. (See Section 5B).

A. COUPLED MOTION (location 100)

To solve the differential equation of coupled motion, the main program calls on a fourth order Runge-Kutta routine[3]. The angles ϕ and ψ, the total escape wheel angle, the angular velocities $\dot{\phi}$ and $\dot{\psi}$, as well as the control parameter g are computed for each time increment. The program continues coupled motion under the following circumstances:

(1) As long as g<0. (Because of the nature of the coordinate system, g is always negative while the pallet pin can make contact with the escape wheel tooth.)

(2) As long as, for a positive (CCW) rotation of the escape wheel, a succeeding absolute value of $\dot{\psi}$ is larger than the one obtained from the preceding computation. This condition is necessary, since in coupled motion when $\dot{\phi}$ is positive, the escape wheel can only drive the pallet but not slow it down. If such a slowdown is indicated, it means that pallet and escape wheel have separated and that free motion is taking place. A similar control is provided for a negative rotation of the escape wheel, which may occur after impact. When coupled motion is terminated, control is shifted to the subroutine containing the free motion equations (location 200). This is done directly if g<0. In case that g≥0, the main program must ascertain whether the preceding computations have been made for entrance or exit conditions, and accordingly, on which side the next contact will occur. In the sample mechanism, g=0 when ϕ is approximately 146° at entrance and approximately 207° at exit. Thus, if g ≥ 0 and ϕ≤150°, all possibility for entrance contact is ended and ϕ must be incremented by the tooth angle δ (see Figures 2 and 5) while ψ must be incremented by the angle $2\pi - \lambda$. For g≥ 0 and ϕ>150°, entrance action follows exit action and ϕ must be decremented by the angle 2δ (see Figure 4, where the new top tooth No. 3 comes into action). At the same time, the pallet angle is decremented by $-2\pi+\lambda$.

B. FREE MOTION (location 200)

After transferring the appropriate initial values from the main program, the subroutine containing the free motion

---

[3]RKGS Routine, IBM, System/360 Scientific Subroutine Package, (360A-CM-0X3) Version III.

7

equations computes the subsequent positions and angular velocities of pallet and escape wheel. In addition to the above quantities, the total escape wheel angle is continually computed. The decision whether or not to remain in this subroutine is made with the help of the sensing parameters f and g'.

If $f>0$ and $g'\leq0$, free motion is continued without indexing. If $f>0$ and $g'>0$, free motion is also continued. Since now contact is no longer possible, indexing takes place.

If $f\leq0$, control is returned unconditionally to the main program. If it finds in addition that $g'\geq0$, indexing takes place and control is given back to the free motion subroutine. In case that both $f\leq0$ and $g'<0$ contact is about to take place (or actually has just occurred). The program must now decide whether this contact just represents a close approach which will be followed by further free motion, or whether it is the beginning of coupled motion. To this end, the quantities V_p and V_s, which stand for the velocities normal to the pallet and escape wheel contact points, respectively, are computed in the entrance and exit free motion tests[4]. The first three cases of of the entrance free motion test of the main program are illustrated by Figure 7. With both angular velocities $\dot{\phi}$ and $\dot{\psi}$ positive, the following three possibilities exist:

(1) If $|V_p|>|V_s|$, the contacting surfaces will separate again and free motion will occur. Control remains with the free motion subroutine (location 200).

(2) If $|V_p|=|V_s|$, the escape wheel will start driving the pallet in coupled motion, and control must be transferred to the coupled motion subroutine (location 100).

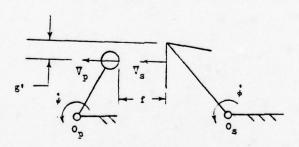Figure 7 Entrance Free Motion Test
($\dot{\phi}$ and $\dot{\psi}$ are positive and distance f is exaggerated)

---

[4]Under the present circumstances, if $\phi<150°$, only top contact can follow, while $\phi>150°$ means that bottom contact will occur.

(3) If $|V_p| < |V_s|$, impact will occur, and control must
be given to the impact subroutine (location 300).

The remainder of the free motion tests are constructed along
similar lines for different combinations of angular velocity directions.

C. IMPACT (location 300)

The subroutine containing the impact equations uses
the current values of the angular velocities $\dot{\phi}_i$ and $\dot{\psi}_i$, and computes the
post-impact angular velocities $\dot{\phi}_f$ and $\dot{\psi}_f$ according to classical rigid
body impact theory. (The tangential impact due to friction has been
neglected).

After impact, control is returned to the main program, and it
must be decided whether free or coupled motion follows. This is
accomplished by comparing the post-impact velocities V_p and V_s in the
impact tests. These are similar to the free motion tests.

If the contact velocities are vectorially equal to each other, or
if the absolute value of the difference of their absolute magnitudes
is less than 2.0 in/sec (considered a small quantity), control is
transferred to coupled motion. If this velocity difference is greater
than the above criteria, computation is transferred to free motion.

5. EXAMPLE MECHANISM

The pin pallet escapement of the M525 fuze was used as
an example mechanism. The following first gives the dimensions of the
basic escapement (standard configuration) and then discusses certain
other data and computed values which are of importance for the com-
puter simulation.

A. DIMENSIONS OF STANDARD CONFIGURATION (See Fig. 5)

The standard configuration has the following
dimensions:

a = .193 in. (mean center distance) b = .1583 in.
c = .0968 in r = .0136 in.
α = 40° λ = 109.337°
I_p = .91 X 10^{-7} lb-sec^2-in.(moment of inertia of pallet)
I_s = .17 X 10^{-7} lb-sec^2-in.(moment of inertia of escape
wheel)

9

B. GEAR TRAIN

The escape wheel of the M525 fuze is driven by a clock spring through a step-up gear train of ratio 45.98. The timing function of the fuze, which involves a delay of between two and four seconds, is accomplished once the spring has rotated the input gear through 310° and with that, the escape wheel through 45.98 times as many degrees. Since the beat of the escapement is generally well stabilized after one tooth cycle, the total fuze time may be obtained by multiplying the time corresponding to 310° of escape wheel rotation by the above gear ratio.

C. STANDARD TORQUE USED IN THE SIMULATION

Measurements on actual fuzes showed that the initial torque on the escape wheels varied between .0177 and .031 in-lb. Since the angle of rotation of the input gear is small, the decrease in torque is also relatively small. Therefore, a constant torque was assumed in the simulation, and its standard value was chosen to be .0177 in-lb.

D. OTHER DIMENSIONS ASSOCIATED WITH STANDARD CONFIGURATION

For purposes of control in the computer program the following dimensions are of interest.

The maximum magnitude for the dimension g, associated with coupled motion is

$$g_{MAX} = -.055 \text{ in.}$$

This occurs for the entrance condition when

$$\phi_{gmn} = 132.4°$$

Because of this value, the program is started in coupled motion for $\phi = 135°$. For exit action, this angle becomes:

$$\phi_{gmx} = 187°$$

The escape wheel angle corresponding to g=0 is also of interest. For entrance action its magnitude is given by

$$\phi_{gon} = 146.3°$$

This angle is responsible for the somewhat larger indexing criterion

LOWEN & *TEPPER

of 150°. (See Section 4.) For exit action, this angle becomes:

$$\phi_{gox} = 206.512°$$

6. INFLUENCE OF VARIOUS PARAMETER CHANGES ON THE TOTAL FUZE TIME

The following reports on the results of numerous computer runs in which a single input or geometric parameter was varied in order to determine its influence on the total fuze time. In all cases, $\mu = .3$ and $\varepsilon = .25$.
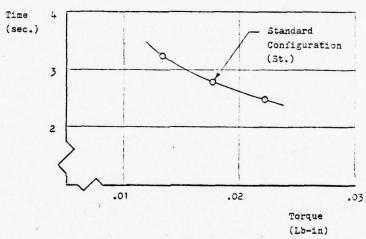
Figure 8   Influence of Escape Wheel Torque on
Fuze Delay Time

Figure 8 shows the influence of the escape wheel torque. One may compare these timing results with those obtained from the well known empirical expression:

$$t_2 = t_1 \sqrt{\frac{T_1}{T_2}}$$

If t_1 and T_1 represent fuze time and torque associated with the standard configuration, one obtains from the above:

For $T_2 = 75T_1$: $t_2 = 2.79\sqrt{\dfrac{1}{.75}} = 3.22$ seconds

and for $T_2 = 1.25T_1$: $t_2 = 2.79\sqrt{\dfrac{1}{1.25}} = 2.49$ seconds

The results of the simulation show excellent agreement with this empirical relationship, which has been confirmed time and again by experiment.

11

B. INFLUENCE OF PALLET MOMENT OF INERTIA
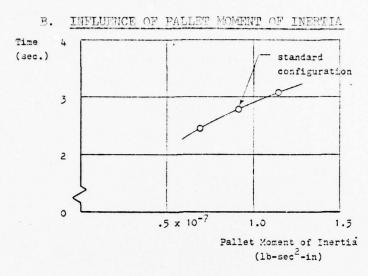
Figure 9 Influence of Pallet Moment of Inertia on
Fuze Delay Time

Figure 9 shows that the total time of the fuze increases with an in-
crease of the pallet moment of inertia. The ratio of any two periods
is approximately proportional to the ratio of the square roots of the
associated pallet inertias.

These results are also confirmed by many reports on experimen-
tation.

C. INFLUENCE OF PALLET-ESCAPE WHEEL CENTER DISTANCE
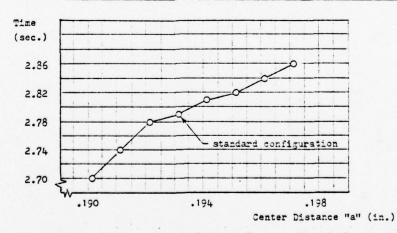
Figure 10 Influence of Pallet-Escape Wheel Center
Distance on Fuze Delay Time

12

Figure 10 indicates that the fuze time increases as the center distance "a" is increased. For a total increase of .007 inches the time increase is approximately 7%. This result is generally confirmed by the experimentation of [2].

D. INFLUENCE OF PALLET RADIUS

Figure 11 Influence of Pallet Radius on Fuze Delay Time

Figure 11 shows a continuous and quite dramatic decrease in fuze time as the pallet radius "c" is varied through .008 inches. Experimentation in [2] gives a good correlation with this result of the simulation.
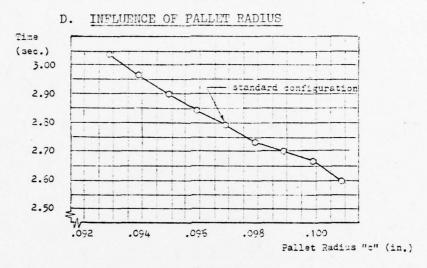
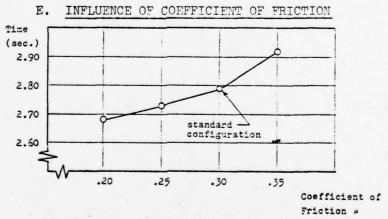E. INFLUENCE OF COEFFICIENT OF FRICTION

Figure 12 Influence of Coefficient of Friction of Coupled Motion on Fuze Delay Time

Figure 12 indicates that the fuze time increases as the coefficient of friction is increased. One would expect that an increase of energy dissipation will slow the mechanism.

F. INFLUENCE OF COEFFICIENT OF RESTITUTION

According to Figure 13, the fuze time increases considerably as the coefficient of restitution is varied from 0 to .5. Reference [1] shows that for $\varepsilon = 0$, coupled motion follows immediately after impact. When $\varepsilon = .5$, entrance action consists of four impacts, three of which are followed by free motion while the fourth is followed by coupled motion. Exit action shows two impacts with the last one followed by coupled motion. Each of the impacts is followed by escape wheel reversal. These multiple impacts and associated motion reversals seem to account for the observed increase in fuze time.

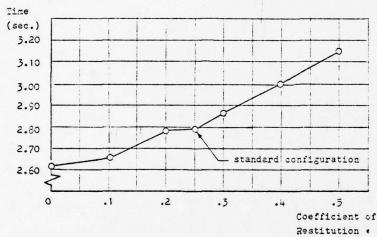Figure.13 Influence of Coefficient of Restitution on Fuze Delay Time

7. REFERENCES

1. G. G. Lowen, and F. R. Tepper, "Dynamics of the Pin Pallet Runaway Escapement", in preparation as an ARRADCOM Technical Report.

2. M. E. Anderson and S. L. Redmond, "Runaway (Verge) Escapement Analysis and Guide for Designing Fuze Escapements," NWCCL TP 860, December 1969, Naval Weapons Center Corona Laboratory, Corona, CA 91720.

ESTABLISHING HABITABILITY FACTORS FOR THE DESIGN
OF OFFICE ENVIRONMENTS

DR. CHARLES C. LOZAR, Habitability Team
U.S. Army Construction Engineering Research Laboratory
Champaign, IL   61820

Purpose

The purpose of this presentation is to document an overall methodology which incorporates experimental design considerations from the social sciences, specifically environmental psychology, and transfers that technology to planning and design application to improve habitability in office environments. The importance of this application is that the habitability factors which are involved in most office environments do not have a firm basis in basic research, and are not well documented in terms of guidance information for designers. This paper will present a discussion of a means of derivation for habitability factors in a particular context of office environments. However, the same methodology will be shown to be applicable to other types of environments, with the process being beneficial to the generation of new basic research, application of new concepts, and continuing accumulation of new knowledge in the area of habitability factors for any environment.

Background

There are four important considerations which point to the need for a methodology to develop habitability factors for environmental research, design, or application. These four considerations are:

1. For a period of time now there has been no sophisticated, well documented basic research from environmental psychology in the determination of habitability factors which affect individuals, either in their office work stations or any other kinds of living environments. Of course, there have been human factor studies, productivity studies, and general evaluations of management climate.(1) However,

consideration for the physical environment as it affects individual's attitudes, preferences, and feelings has not developed to the point where it can provide design decision guidance, or technology transfer of basic research.

2. This basic technology transfer, information, lacking in the research literature, is precisely the kind designers need to apply to physical environments. In view of the lack of this kind of information, most schools of design operate on an intuitive basis, predominated by the master studio student relationship. Any guidance received is based on intuition and experience and is not verified empirically through research data or documented in a way in which it can be used over and over again.

3. It follows logically from the first two considerations that there is a need to improve the overall research information quality used to design physical environments, in this case offices. However, because of the limitations of the first two items, it is impossible to document research knowledge which can accumulate and thereby provide a basis for using each design of a new office as a field test to gain new habitability information. Consequently, there is generally no improvement in the quality or depth of research information available to address the problem of habitability factors as they apply to specific environments.

4. The information from environmental psychology that presently exists in this context, is then mostly academic and does not seriously attempt to make effective technology transfer since it does not provide a means of determining either the return-on-investment payoffs or guidance in the allocation of funds based upon the determination of habitability factors.

Therefore, although there are certain factors such as privacy, professional image, room occupancy, etc. which appear in the general literature on habitability,(2) there is little consideration of a means of transferring this information to application in the field and, thereby, providing a means of technology transfer and return-on-investment. Furthermore, the existing research literature suffers from serious methodological deficits in terms of its application to field situations.

Previous Research and Its Limitations

General and intuitive descriptions of office work environments exist in the commercial literature. However, there are also some recent studies from environmental psychology which suggest that the office environment is a viable area for empirical research and investigation. These studies can be divided into two general groups: those which provide methodologically tight organization of variables

and assume some scientific generalizability, and those which take
behavioral and psychological information and translate it into pro-
gramming recommendations for design planning. Examples of the former
are comparisons of open and closed plan offices by Brooks and Kaplan
(3) and Nemecek and Grandjean. (4) Results of these investigations
suggest major important habitability variables in office settings.
Examples of studies which translate user need research into design and
planning recommendations are those described by Deasey, (5) Moleski,
(6) and Davis. (7) However, in order to explore some of the more
relevant habitability relationships suggested in the earlier discus-
sion, we need to identify specific major environmental issues which
may have an impact upon employee performance or productivity. Sug-
gestions for the purposes of examining previous research are profes-
sional image, privacy, room occupancy, office partition type, win-
dows and views, and color.

Professional image is both an organizational and a personal
variable. If an office work station looks professional, well kept, and
neat, some qualitative things about the pride of the employee are
suggested. This form of a sense of identity is discussed by Steele
(8) as a reflection of what the organization "is." In another dis-
cussion of identity, pride, and self-image, Braum (9) suggests that
self-image and individual worth may be related to the physical en-
vironment and productivity. In general, the image of the environment
and its relationship to the person's self-image is not a well-
researched area.

Privacy in offices is a very topical issue as many large organiza-
tions shift to open-landscaped planning arrangements. Altman (10)
suggests that privacy is selective control of access to one's self from
the group. In offices, changes in privacy can occur easily since by
varying the partition types, one has a great deal of control over
access to other persons. Johnson (4) suggests that control over pri-
vacy behavior is a means of reducing some forms of stress and also a
way of governing a path towards one's work objectives. Implications
for office environments are obvious, but the specific question is how
is one to manipulate all of the interacting variables in the physical
and social environment to achieve this? This is, in particular, where
the programming efforts of the designer supplement the statistics of
the psychologist.

Room occupancy and partition use in office design is related to
concepts of population density and personal territory, both perhaps
related to privacy. Density, with research into its effect upon be-
havior, has been discussed by Calhoun, (12) Sommer, (13) and
Rappoport. (14) It is worth noting that this subject is still under
considerable controversy. Some researchers would suggest an optimum

density in an office space is a concept relating to the nature of the work which generates some level of noise which is transmitted over some physical distance determinant of density. Partition type, that is, the mode of division of office work spaces, is a little researched area. In a study by Dinnat and Gibbs, (15) comparisons were made between open-plan, bank-screen, and landscaped partition types. Also, human factors studies by Probst (16) in the development of an office product line have provided much useful data about behavior-work station interaction and have been used in designing work station components, such as partitions. Layout of partition type is an ambiguous area of research. In studies comparing office arrangements, Wells (17) noted small office arrangements were more conducive to small group identity and open plans were more conducive to interpersonal contact. Brookes and Kaplan (18) noted that landscaped offices did not function as well as conventional offices, but were aesthetically more pleasing. However, Zeitlin (19) notes that landscaped offices tend to reduce privacy. Partition type then is a determinant of the mode of layout, and in turn can influence attitudes toward habitability factors in the whole office.

Windows and views and color are important to all workers. How important is, however, an unclear area in the research literature. In a summary of research dealing with psychological reaction to environments with and without windows done by National Bureau of Standards, (20) the results were inconclusive. That is, there seemed to be general disagreement on the importance of windows to employee performance throughout most studies. Color in offices and work areas is another important issue in office planning, but still subject to much intuitive decision making. In most office studies, the question of color preference is asked of the employees, but documentation of effects of color is hard to find. A study by Goodfellow and Smith (21) even suggests that color of rooms has no effect upon psychomotor task accomplishment at all.

In summary then, much of the research in this realm of habitability factors and office design is somewhat formative, inconclusive, and exploratory. That is, the relationships between physical parameters and social effects are not clearly defined so that the empirical inspection of relationships can occur for planning purposes. Furthermore, there are serious limitations relating to the generalizability of information which hinder the application of previous studies to the development of well-based habitability factors. These limitations are as follows:

1. Most of the academic studies are one-shot affairs and do not use a standardized questionnaire over and over again in various field tests. Therefore, although the questionnaire may be pretested for

local population, its generalizability is limited. Consequently, the comparison between different field environments in most of the previous literature cannot occur, and generalizability is limited or not possible at all. (22)

2. There is a differentiation of experimental methods in all of the research literature on habitability parameters. There are almost no studies which are replications; and, generally, the literature does not contain enough information on methodology to allow a valid replication. Therefore, generalizability, without experimental bias due to different field methods, is impossible. Since the methodologies and generally the questionnaire instruments are not comparable, there has been no effort to compare data across field studies to verify a general basis for habitability parameters. (23)

3. Therefore, the environmental psychology research literature presently in existence appears fragmented, has many different factors involved, and does not provide a consistent basis for overall generalizability to habitability factors. This situation inhibits the development of guidance and the achievement of technology transfer.

4. Finally, there is little emphasis on the allocation of funds in a rational, acceptable way to determine returns on design investments related to habitability factors. Since this data does not exist, most of the dollar allocations in any new building or renovation of an office tend to still be done intuitively, and consequently there is no possibility of doing an estimate of a return-on-investment.

In summary then, there is the lack of a consistent experimental method which will provide a long-term basis for deriving generalizable habitability factors. In view of these deficits in the research literature, a process has been developed and documented herein to rectify this situation. This process enables both the designer and the researcher to work together to build a cumulative process of deriving habitability factors from various office settings, each one regarded as an individual field study contributing to both theory building and technology transfer.

Field Studies to Create a Data Base

The purpose of conducting field studies is to provide an experimental basis for determining habitability factors. In administrative facilities such as offices, data from users can be used to establish factors of habitability, such as privacy, space, view, noise, or image. Evaluation of these environmental factors can then be applied to interior design solutions in order to improve the habitability for all office occupants. The methodology of this type of field research is to: 1) analyze results from a before- and after-renovation study, and 2) generate design criteria for layout and workstations of offices.

LOZAR*

The approach to determining habitability factors for office occu-
pants consists of the design of a before-after experiment in which
certain parameters of the environment (such as floor space, distance
to next person, and degree of enclosure of the work-station) can be
measured. For example, an existing office building of 130 persons
participated over a one year period in an evaluation. The initial
comprehensive survey of occupant attitudes and behaviors indicated
certain environmental conditions related to the habitability for the
office occupants. New office layouts and workstation arrangements
were designed and installed. The workstation components were designed
in such a way that a within-groups experimental design for some para-
meters (high vs. low partitions, floor area variations, etc.) was
possible.

After an occupancy period of six months in the new office environ-
ment, the users were again surveyed. Data analysis consisted of be-
fore-after comparisons, satisfaction with privacy, space, image, noise,
etc. and with individual aspects of the workstation such as floor area,
storage, work surface, etc. Factors of habitability (such as work-
station image, privacy, and furniture satisfaction) were further ana-
lyzed with regressions to indicate shifts in user's cognitive aware-
ness of the environment in the before and after office conditions.

The data from this office study is then incorporated into a data
base using Statistical Package for Social Sciences programs (SPSS)
(24). Other office contexts can then be added to this original study
in the data base to provide a basis for comparison. These various
contexts have included the Army Research Office, the Construction
Engineering Research Laboratory, the National Aviation Facilities
Experimental Center, Foreign Science and Technology Building, etc. for
a total of over 700 individual subjects.

To provide for a basis for generalizability in a habitability
data base for offices, there must be three common elements across all
office research contexts. They are as follows:

1. Most of the office environments in the before condition are
similar in character, arrangement, floor area, lighting levels, etc.
Since most of these office environments are from Government technical
installations, the overall organizational character of the sponsoring
organization is somewhat the same also.

2. A second commonality amongst the environments studied is the
demographics of the subjects. Most individuals in these offices were
of three groups: (1) managers, (2) research/technical-oriented in-
vestigators, or (3) support personnel such as secretaries, etc.

3. The same commonality across groups exists in terms of their tasks; that is, subjects are either involved in research generation and investigation, support of that research endeavor, or management of that research or technical endeavor.

Therefore, it is reasonable to assume that there is some degree of comparability across the office environments in the data base. The demographics, the tasks, the environments, and the settings are similar. The experimental prerequisites are sufficient for the creation of a comparative data base. In summary, then, the information from the before and after evaluations of the field studies are structured into an SPSS data base, and the data base used to investigate habitability factors, such as privacy, within the office context being evaluated, and comparatively over other office contexts. A simple matrix representation of the data base is shown in Fig. 1.

Methodology for Empirical Habitability Factors

By careful statistical manipulation of the data base, the experimenter can isolate the individual effects of separate, independent variables operating simultaneously on a single dependent variable. By doing this comparatively across a number of individual settings, the researcher can develop a basis for theory building to establish habitability factors in offices. A major goal of theory building is to make use of the deductive types of arguments which can go beyond common sense but still can be empirically verified. Therefore, it is necessary for the experimenter to be cognizant of the limitations of his experimental design and still be perceptive of inferences which may be made by comparing individual designs. In most experimental designs referenced in this paper, Campbell and Stanley's definition of the pre-test/post-test control group design is most applicable. (25) In some instances, in determining the effect of complex research variables on one another, a Solomon four-group design is used. In conditions where it was impossible to do a post-test, a pre-test or static group comparison was used.

In order to generate habitability information in the form of non-generalizable guidance, standard statistical analyses are used, which would result in a statement of guidance as shown in Fig. 2. This could be used for the particular office design, but would not have generalizable validity until comparatively checked with other office evaluation results in the data base. To make the information more generalizable, the researcher determines a habitability factor and develops statistical comparisons as shown in Fig. 3.

By constructing a data base of this nature dealing with over 1200 (before and after conditions) individual subjects at five different office sites with similar demographics, environments, tasks, and

DEMOGRAPHICS

AGE
SEX
EDUCATION
JOB TYPE
FUNCTIONS
TENURE
RANK
ETC.

FLOOR AREA
NOISE LEVEL
LIGHTING LEVEL
SIZE OF WINDOW/VIEW
% FLOOR IN FURNISHING
OTHER PHYS. PARAMETERS

STIMULUS LEVEL OF ENVIRONMENT

FURNITURE   WORK STATION   ROOM   BLDG COMPONENT   TOTAL BUILDING
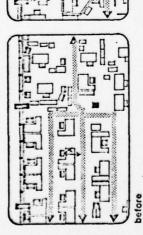
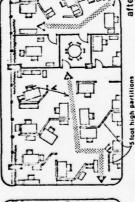Fig. 1: Matrix Representation of Office Facility Data Base

This diagram shows generic data collected and structured for each office study. Data can be analyzed for individual offices or for all offices combined.

Problem Statement: Many distractions occur in open offices because of the passage of persons near a workstation. These distractions can be modified by the extent and location of circulation paths.

Requirements: Office workers have a requirement for privacy from circulation paths in open office areas.

Criteria: No specific criteria exist dealing with the relationships between circulation and lack of privacy (as distractions from noise and movement).

before

after

5 foot high partitions

Research Commentary: In the before condition, the major circulation paths were directly through the open work areas of exposed workstations. This was changed in the after condition so that most circulation was in the hallways and the workstations were "enclosed" with 5-foot high partitions. The distance from a circulation path was used as a design variable, thus respondents were divided into four groups: (1) those three feet away from the circulation path, (2) those four to six feet, (3) seven to nine feet, and (4) ten or more feet as measured on the plans. Their responses to items dealing with (1) conversations as a disturbance, (2) people coming into their areas, (3) visual distractions, (4) control of privacy, and (5) perception of adequate privacy were analyzed.

The results presented in the tables below indicate "general" improvement across all variables in the after (partitioned) condition. However, the distance of 10+ feet away is consistently the category of major improvement. It would appear the 10 feet distance is a threshold where the visual and acoustical attenuation components (of the after condition) accomplish their design purpose.

Q. 88: Conversations in my room disturb my ability to concentrate.

distance from circulation path	% agreeing		
	before renovation	after renovation	improvement
-3 ft.	90	75	+15
4-6 ft.	57	83	-26
7-9 ft.	92	76	+16
10+ ft.	86	52	+34

Q. 92: I have a high degree of control over my privacy in my room.

distance from circulation path	% agreeing		
	before	after	improvement
-3 ft. %	0	13	+13
4-6 ft. %	0	0	0
7-9 ft. %	15	23	+8
10+ ft. %	7	32	+25

Q. 21: The privacy I now have is adequate for my tasks.

Q. 90: People keep coming into my room and disturbing re.

distance from circulation path	% agreeing		
	before renovation	after renovation	improvement
-3 ft.	91	62	+29
4-6 ft.	57	71	-14
7-9 ft.	69	46	+23
10+ ft.	77	24	+53

Q. 93: I have many visual distractions in my office which are disturbing.

distance from circulation path	% agreeing		
	before	after	improvement
-3 ft. %	48	25	+23
4-6 ft. %	43	28	+15
7-9 ft. %	17	29	-12
10+ ft. %	57	18	+39

distance from circulation path	% agreeing		
	before	after	improvement
-3 ft. %	10	31	+21
4-6 ft. %	7	17	+10
7-9 ft. %	23	35	+12
10+ ft. %	7	45	+38

Guidance: It has always been good design practice to keep circulation paths away from work areas. In open office planning, this is a difficult since almost any path between two workstations is a reasonable path for circulation. Since privacy is impacted by the proximity to circulation paths, the circulation paths should be short dead-ends that are screened from the actual workstations, with any major circulation at least 10 feet away.

Fig. 2: Development of Guidance for Habitability Factor
Chart shows result of analysis of a single parameter
for one office context.

Satisfaction with Furnishing = Constant + b_1 comfortable furniture + b_2 modern furniture + b_3 sturdy furniture + b_4 high quality furniture

	Constant	b_1 comfortable furniture	b_2 modern furniture	b_3 sturdy furniture	b_4 high quality furniture
Office A Movable Partitions (n = 23) =	-.628	.278	.670	.273	-.021
Office B Open Office Area (N = 91) =	.308	.435	.175	-.013	.360
Office B High Partitions (N = 23) =	.426	.492	-.129	.346	.246
Office C Bank Screen Partitions (N = 23) =	1.28	.409	-.230	.255	.585
Office D Private Offices (N = 59) =	.623	.186	.314	-.108	.440
Office D Open Office Area (N = 339) =	.506	.255	.300	.065	.372

Fig. 3: Comparative Regressions on Satisfaction with Furnishings
Results from comparison of a single habitability parameter over office A, B, C, and D. Inspection of shifts in coefficients indicate different attitude responses to office conditions. Most stable factors are comfortable furniture and high quality furniture.

settings, the earlier criticisms of the lack of an empirical basis for
developing habitability factor research are negated. By using the
data base, it is possible to determine, recognizing some inherent
methodological biases, generalizability across various factors and
groups. An example of this process follows.

Establishing Habitability Factors

One basis for determining habitability factors is to start with
the evaluation of the validity of certain hypotheses about the inter-
action of man, his attitudes, and the environment in the office. These
hypotheses are generally built out of the research data which exists in
the literature. For instance, the hypothesis that there is a relation-
ship between the distance from workstation to the circulation paths
and ratings of overall privacy in an office is documented in a number
of publications. Using the data base, these hypotheses can be tested
for single office studies and comparatively over other studies. An
example of this type of analysis is presented in Fig. 2. Here, simple
statistical analysis is used to indicate trends. This is adequate for
a single office context, but has little generalizability. By compar-
ing results over different office studies, the researcher can feel
more confidence in his interpretation of data.

Taking as an example, a simple regression analysis to determine
the major components of satisfaction with furnishings, the results are
seen in Fig. 3. These results are for five office environments and
would be generalizable to other office environments. The major fac-
tors which contribute most to the variance of the dependent variable
are shown underlined in this table. Although this is not a complete
analysis, it does give the reader an indication of how a designer and
a researcher would collaborate using the data base to determine major
habitability factors and taking that information to determine prior-
ities for renovation scheduling or priorities for the allocation of
dollars to a renovation. In this way the application of the data base
to existing problems would accomplish a number of specific improve-
ments over previous literature. Therefore, the researcher could use
the data base to create a number of regression analyses for comparison
across different physical office conditions to determine the amount of
variance accounted for, and the designer would use these equations as
a guide to the allocation of funds to improve habitability in office
environments. This would enable the designer to make an evaluation
for an existing environment and that evaluation could be generalizable
to a new building design.

In summary then, a methodology is presented for establishing
overall habitability factors with some degree of generalizability.
This is done by creating a standardized methodology which addresses the
limitations of a pre-tested questionnaire, the comparability of data,

and the comparison of research environmental contexts. By using the data from different contexts in a comparative data base, the research-er and the designer are able to collaborate in the establishment and the application of information. Furthermore, each application of the information in a new context represents a step forward in another ex-perimental design as it is added to the existing data base. This new application can then be used to replicate previous studies, to vali-date previous information and hypotheses, or to develop new empirical or intuitive insights into other habitability factors.

The point of this then is that the designer no longer lacks basic habitability research information, but now has a step by step method to improve the overall quality of his design decision making along with the researcher improving the overall quality of his basic re-search in real field contexts.

Developing Guidance for Technology Transfer

This method for determining habitability factors provides a basis for developing guidance to apply research information from the data base to new field situations. In simple terms, this is regarded as technology transfer. By using this approach, that is, standardizing the method across environments and comparing from a data base, one is able to develop three specific technology transfer impacts at dif-ferent levels of application. All of these contribute to the develop-ment, support, and application of habitability factors in office en-vironments, or may be used appropriately to other environments. These are:

1. Information developed from the application of the data base can be translated into "design guides" which are similar to catalogs specifying types of components to go in a new office environment. The specific difference between these and ordinary catalogs is that the elements of the design guide are selected on the basis of the infor-mation derived from an analysis of variables within the office data base. As an example, one might select three different manufacturers of desk types as being acceptable to improving overall satisfaction in office environments based upon a comparative analysis of satisfaction with different furniture types.

2. Information derived from the data base can be used for small problem field consulting. This is a case in which a client is in need of some level of decision making and does not require a major research effort, his variables may be related to those already in the data base. By manipulating the data base, one can determine some level of predict-ability in order to suggest actions for implementation. An example may be a client who wishes to know the impact of energy conservation measures in a particular type of office environment. One can partition

the group of subjects in the data base with low lighting levels and determine the minimal lighting level which will conserve energy and yet provide adequate satisfaction for office occupants.

3. Finally, one of the benefits of the construction of a field-oriented methodological data base dealing in habitability factors is that it provides the basis for project-to-project comparison, validation, reliability testing, and generation of new hypotheses which serve to build the basis for a theory of habitability in office environments. This same methodology for field application and comparability can also be used for housing, hospitals, residences, etc. The application of this procedure could be of great benefit to the entire field of habitability research and provide a systematic approach to defining, verifying, and validating habitability factors where one presently does not exist.

REFERENCES

(1) Proshansky, Ittleson, Rivlin (ed.), Environmental Psychology: (1st Edition) Man and His Physical Setting, Holt, Rinehart, and Winston, 1967.

(2) Proshansky, Ittleson, Rivlin (ed.), Environmental Psychology (2nd Edition) Man and His Physical Setting, Holt, Rinehart, and Winston, 1976.

(3) Brookes M., Kaplan, A., The Office Environment: Space Planning and Affective Behavior, Human Factors 1972, 14, p 373-391.

(4) Nemecek, J., and Grandjean, E., Results of an Ergonomic Investigation of Large-space Offices, Human Factors 1973, 15, p 111-124.

(5) Deasey, C. M., Design of Human Affairs, John Wiley and Sons, N.Y., 1974.

(6) Moleski, W., Behavioral Analysis and Environmental Programming for Offices, in Lang et. al (ED) Designing for Human Behavior, Dowden, Hutchinson, and Ross, 1974 p 302-314.

(7) Davis, G., Applying a Planned Design Process and Specific Research to the Planning of Offices, in Carson (ED) Man-Environment Interactions, Dowden, Hutchinson, Ross, 1974, Vol II, p 63-90.

(8) Steele, Fred, Architecture and Organizational Health, in Archea and Eastman (ED), Proceedings of Second Annual Environmental Design Research Conf., Halstead Press, 1970, p 252.

(9) Braum, "Identity Behavior and Architectural Form," unpublished paper, Dept. of Architecture, University of Illinois, 1964.

LOZAR*

(10) Altman, I., _The Environment and Social Behavior_, Brooks/Cole, Monterey, Cal., 1975, p 52-58.

(11) Johnson, S., unpublished communication with author, June 1974.

(12) Calhoun, J. B., Population Density and Social Pathology, _Scientific American_, 1962, p 139-145.

(13) Sommer, R., Studies in Personnel Space, _Sociometry_, 1959, 22, p 247-260.

(14) Rappoport, A., Toward a Redefinition of Density, _Environment and Behavior_, Sage Publications, Vol 7, No. 2 June 1975, p 131.

(15) Dinnat, R. M., Gibbs, W., _Cost-Effectiveness of Three Different Interior Open-Type Offices_, Technical Report D-2, U. S. Army Construction Engineering Research Laboratory, 1973.

(16) Probst, R., _The Office: A Facility Based on Change_, Herman Miller Research Corp., 1968.

(17) Wells, B., The Psycho-social Influence of Building Environment: Sociometric Findings in Large and Small Office Spaces, in R. Gutman (Ed.) _People and Buildings_, Basic Books, N.Y.

(18) Brookes and Kaplan, Ibid.

(19) Zeitlin (This is not a report but part of a study referenced in Brookes and Kaplan [1]).

(20) U. S. Department of Commerce, _Windows and People: a Literature Survey_, NBS Building Science Series #70.

(21) Goodfellow, R., Smith, P., Effects of Environmental Color on Two Psychomotor Tasks, _Perceptual Motor Skills_, 1973, 37, p 296-298.

(22) See _Proceedings_ of the Environmental Design Research Association Conferences, 1969-1977 (various publishers).

(23) Lozar, C., "Measurement Techniques Toward a Measurement Technology," in Carson, D., _Methods and Measures, Man-Environment Interactions--The State of Art in Environmental Design Research_, Milwaukee, Wis., p 185, 1974.

(24) Nie, Hull, et. al., _Statistical Package for the Social Sciences_, McGraw-Hill Books, 1975.

(25) Campbell and Stanley, _Experimental and Quasi-Experimental Designs for Research_, Rand-McNally, 1963.

THE EFFECTS OF PHYSICAL AND CHEMICAL
PROCESSES ON TWO-PHASE DETONATIONS

PAI-LIEN LU PHD, NORMAN SLAGG PHD, BARRY FISHBURN PHD
US ARMY ARMAMENT RESEARCH & DEVELOPMENT COMMAND
DOVER, NJ   07801

INTRODUCTION: During the past fifteen years, many
studies have been conducted on the shock ignition and propagation of
detonation in two-phase mixtures consisting of liquid fuel drops with a
gaseous oxidizer. In most cases, oxygen-rich atmospheres and pure
fuels were employed. Physical mechanisms which describe the energy
transfer in a two-phase detonation were suggested and generally accept-
ed. The sequence of processes that occur following the passage of the
leading shockwave of the detonation is: (1) breakup of fuel drops via
boundary layer stripping; (2) ignition of the fuel micro-mist that
forms in the wake of individual drops and (3) energy transfer to the
shock front via local blast waves originating from individual burning
drops. However, in general practice, two-phase detonations do not occur
in oxygen-rich atmospheres and often involve fuels containing compounds
which modify the combustion chemistry. Lower oxygen concentration and
fuel additives were expected to significantly change the physical
mechanisms outlined by the work on pure fuels in oxygen-rich atmo-
spheres.

The effects of chemical sensitizers on the ignitability and
detonability of gaseous mixtures have been investigated for many years.
But the effects of the additives in a two-phase mixture, where fuel-
drop breakup is involved, were unknown. This study was therefore,
initiated to determine the extent to which additives affect the ignit-
ability and detonability of a two-phase mixture.

The study provides a complete examination of the role of
chemical additives on fuel detonability. First, detailed laboratory

experiments were carried out on shock ignition of individual fuel drops and on detonations propagating in monodisperse fuel sprays under controlled conditions. Then, large scale field tests of explosively-disseminated fuel aerosols were carried out, to demonstrate that the effects of additives, manifested in the laboratory, occur in the field. Each step in this sequence will be described individually.

EXPERIMENTAL PROCEDURE

A.   Single Drop Studies

A combustion driver was developed and used to achieve a high enough shock strength to ignite hydrocarbons in air. The fuel drop was suspended on the horizontal section of a hypodermic needle in a shock tube test section. Shock velocities were obtained by measuring the time required for the shock to pass between two pressure transducers that were flush mounted on the inside wall of the test section. Light emission from the fuel drop was monitored by a photodiode whose output was displayed on an oscilloscope. Ignition delay was defined as the time interval between passage of the shock wave over the drop and onset of the light signal.

B.   Detonation Studies

The detonation study was conducted in a vertical detonation tube apparatus consisting of drivers, a fuel drop generator, the detonation tube and dump tanks. A schematic of its layout is shown in Figure 1. The stainless steel detonation tube is 457 cm long and has an inner cross section of 4.1 x 4.1 cm. Two driver sections are mounted opposing each other at the top of the tube in such a way that detonation of the driver gas forms two blast waves which collide along the tube center line and initiate the fuel drop-air mixture.   A stoichiometric hydrogen-oxygen mixture with 689 kPa initial pressure was used as the driver gas.

Two different fuel drop generators were used. One is a vibrating hypodermic needle type used to generate monodisperse aerosols of 700 micrometer or larger drops. With this generator, the fuel-air ratio was controlled by changing both the drop size and the number density of fuel drops generated in the tube. The other generator is an ultrasonic nebulizer which generated a mist of fuel drops ranging in size from 0.5 to 10 micrometers. With this generator, the fuel-air ratio was controlled by diluting the very rich fuel-air mixture from the nebulizer with various amounts of secondary air.
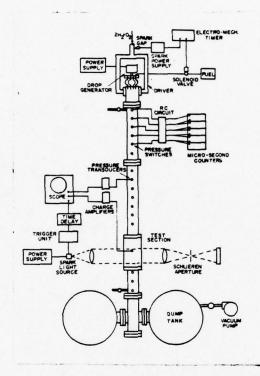
Figure 1. Schematic Diagram of Two-Phase Detonation Apparatus.

Shock velocity in the tube was obtained by measuring the time required for the shock to pass between pressure switches along the tube. Pressure-time profiles at various locations along the tube were measured by pressure transducers flush mounted in the tube wall. A spark schlieren photographic technique was used to record the detonation flow field in the tube.

C. Field Test

The field tests were conducted at the Naval Weapons Center, China Lake, California. Fuel cannisters were made of two gallon plastic jars with holes in the caps for insertion of central burster charges which disseminate the fuel. Central bursters were made from 1.27 cm internal diameter by 22.9 cm long tubes, each with a welded flange at one end which seals to the inside of the cap with an O-ring. The burster charge inserted in the burster tube was a 1.27 cm diameter by

3

17.8 cm long rod of PBXN-201 explosive. Initiation of a fuel-air cloud was accomplished by a small charge of N-5 explosive supported on top of a 122 cm tall wooden stand located 183 cm from a fuel cannister. All canisters were suspended 122 cm from the ground. In this study, the initiators were fired 110 milliseconds after the burster charge was fired, which was long enough to allow the fuel-air clouds to approach their maximum dimensions. Three piezoelectric pressure gages, spaced 152 cm apart, were flush mounted in the ground to measure pressure-time profiles from the detonation. The entire dissemination and detonation process was filmed using high speed cinemotography.

D. Fuels

Heptane, which has some physical properties similar to gasoline and has the poorest "knock" rating in internal combustion engines, was chosen as the basic fuel to be studied. The two additives used to alter chemical properties were butyl nitrite and normal propyl nitrate (NPN). Physical properties of each heptane and additive mixture used in this study are summarized in TAble 1. Dry air was the oxidizer used in all experiments.

Table 1

Physical Properties of Heptane/Additive Mixtures

Additive	Amount of Additive % By Weight	Mixture Density (G/CC)	Mixture Viscosity (CP)	Mixture Kinematic Viscosity (CP) (G/CC)	Mixture Surface Tension (DYNES) (CM)
None	0	.683	.429	.628	23.2
N-Propyl Nitrate (20% by volume)	29	.762	.444	.583	24.5
N-Propyl Nitrate (10% by volume)	14.5	.725	.438	.604	24.5
Butyl Nitrite (10% by volume)	12.9	.711	.448	.630	23.6

4

RESULTS AND DISCUSSION

A. Single Drop Studies

In Figure 2, schlieren photographs are presented of shock ignited heptane drops with various additives in air. In contrast to the previous studies of shocked fuel drops in pure oxygen by us (Ref 1,2) and by Kauffman (Ref 3), there is no local blast wave originating from the wake of each disintegrated fuel drop. This result was not too surprising since the rate of energy release from the micromist formed by mass stripping of the parent drop is expected to be a function of the oxygen concentration. With a high rate of energy release, an exploding or detonating mode of combustion is more likely and local blast waves can be generated. Since the energy generation rate is slower in air, only a deflagrating mode of combustion is apparently possible.
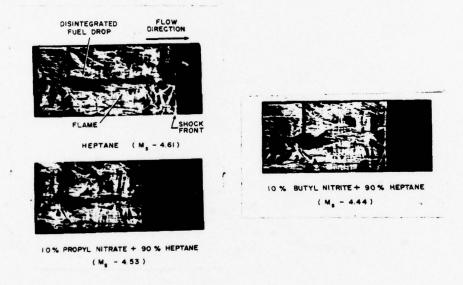
Figure 2. Schlieren Photographs of Shock Ignited Fuel Drops with Chemical Additives in Air.

Figure 3 shows measurements of ignition delay times of heptane and heptane-additive mixtures. The ignition delay time t_{ig} is plotted against the inverse of the free stream air temperature, T_2 , behind the incident shock. Least mean squares plots of the experi-

LU, SLAGG, FISHBURN

mental data are also shown in the same figure. It is seen that both
additives reduce the ignition delay time. A theoretical study by
Fishburn (Ref 4) indicates that the overall breakup time (t) of a drop
in a supersonic gas flow occurs on a time scale given by

$$t = \left(\frac{\rho_{gas}}{\rho_{fuel}}\right)^{\frac{1}{2}} \frac{U_{gas}}{D_{fuel}}$$

where ρ is the density, U_{gas} is the gas velocity and D_{fuel} is the drop
diameter. Thus, the drop breakup rate is not very sensitive to the
physical properties of the fuels. In view of the properties listed in
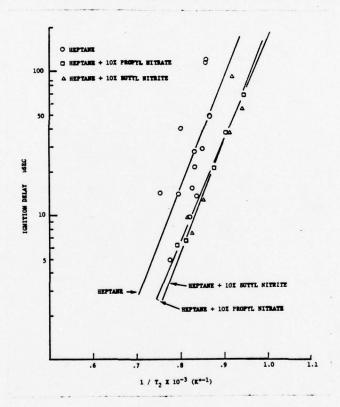
Figure 3. Ignition Delay Time of Shock Ignited Fuel Drops
with Chemical Additives in Air.

6

Table 1 the additives are not expected to significantly effect the breakup time component of the ignition delay time. Thus, the reduction of ignition delay time by these additives must be due to chemical processes. It is known that the $RO-NO$ and $RO-NO_2$ bonds in the nitrites and nitrates are much weaker than the C-C and C-H bonds in hydrocarbons (Ref 5). It is also known that additions as small as 1-2% of NO_2 can reduce the ignition delay of methane - air mixtures by 1/2-1/3 (Ref 6). Though the detailed chemical process is not known in our system, it is believed that the additives act as a source of the free radicals RO, NO NO_2, and NO, which speed initiation of chain reactions resulting in a flame and/or explosion.

B.   Detonation Propagation

The effect of the propyl nitrate and butyl nitrite on the detonability of a heptane drop-air mixture can be seen in the 1400 μm drop mixture wave velocity history vs tube length shown in Figure 4. For comparison purposes, the blast wave velocity in an empty tube is also shown.

The wave speed decays in both heptane, and heptane + 10% normal propyl nitrate (NPN) air mixtures, but more slowly when NPN is present. In all other cases the wave speeds all reach a steady state after a transition region, which indicates the establishment of self-sustained detonations.

It is interesting to speculate on the cause of the differences in detonability of the mixtures mentioned above. In the case of heptane - oxygen mixtures, it is obvious that the pure oxygen atmosphere is a significant factor in the detonability. It has been r reported by Kauffman (Ref 7) that the increase of oxidizer concentration by a factor of 3 will decrease the ignition delay time of shock ignited fuel drops by a factor of 4. It is believed that due to the short ignition delay time, the energy released by the burning fuel drops is better able to couple with and support the incident shock. In the cases of 90% heptane + 10% butyl nitrite and 75% heptane + 25% NPN mixtures in air, the picture is less clear. One possibility is that only the reduction of ignition delay time caused by the addtion of nitrite and nitrate improves the coupling between the shock front and combustion zone. In addition, it is also possible the nitrite and nitrate, through the formation of the RO, NO_2 , and NO radicals, lead to new chain reaction paths which increase the overall rate of energy release from the burning drops. This would improve coupling between the shock wave and combustion zone to an even greater extent. One or both of these possibilities may be occurring.
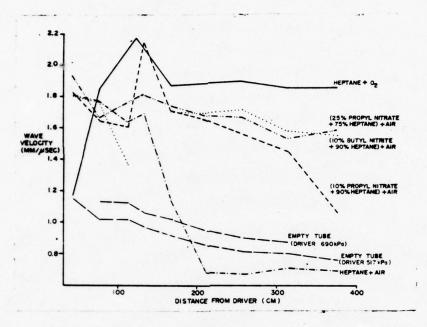
Figure 4.   The Effect of Additives on the 1400 μm Heptane
Drop Mixtures Wave Velocities.

The effect of the fuel drop size on the detonability of
heptane - additive - air mixtures can be seen in Table 2.   Because of
the high vapor-pressure of heptane and its additives, most fuel drops
in the fog mixture probably vaporize prior to detonation.

Table 2

Effects of Additives and Drop Sizes on the Detonability of Heptane-Air Mixtures

Drop Size

Fuel	1400 μm	700 μm	Fog (∿10 μm)
Heptane + 25% NPN	detonation		
Heptane + 20% NPN		detonation	detonation
Heptane + 10% NPN	no detonation	detonation	detonation
Heptane	no detonation	no detonation	detonation

8

These results show both the nitrate additive and small size drops enhance detonability. This is not too surprising, since both the additive and small drop size are suggestive of faster rates of chemical energy release and thus better coupling between the shock front and the combustion zone. It has been shown (Ref 8) in gaseous detonation that the quenching condition of a detonation wave depends on the coupling between shock front and combustion zone. Clearly, the hazard presented by a particular fuel-air mixture can be drastically altered by the addition of small amounts of certain additives or by decreasing the drop size. This means that the danger of any fuel-air mixture should be assessed individually.

Currently various computer codes have been developed to calculate detonation properties for any exothermic system; however, no theoretical techniques have been developed to predict the detonation limits, as kinetics and various energy coupling mechanisms will define the actual limit conditions. Figures 5 and 6 show calculated and experimental detonation wave speed vs equivalence ratio of heptane + 20% NPN and pure heptane mixtures in air. The Tiger code was used to calculate theoretical Chapman-Jouguet wave speeds. Mixture compositions where detonations cannot be sustained in the tube are also indicated in the figures. In general, the experimental wave speed data follow the calculated wave speed fairly well. In the heptane + 20% NPN mixtures, the measured wave speed of 700 μm drop mixtures show a larger decrement from the calculated curve than the measured wave speed in fog mixtures. This is due to larger energy losses to the tube wall in the thicker reaction zone produced by the larger drops. (The reaction zone structure will be described later.) It is seen from Figures 5 and 6, that both small drop size and addition of propyl nitrate expand the fuel rich detonation limit, but have essentially no effect on the fuel lean limit. The 700 μm drop mixture fuel lean limit is shown lower than that of the fog mixtures. The cause of this difference is not clear. One possible explanation is that fuel vapor originating from fuel drops falling down the tube is sufficient to make up the difference in fuel concentration. No account was taken of this vapor in calculating the fuel concentration. Another possibility is that the detonation propagation mechanism is controlled by local conditions in the reaction zone, and the local fuel concentration in the 700 μm drop mixture is quite different from the overall fuel concentration in the fog mixture.

For comparison purposes, the detonation limit of propylene oxide (Ref 9) is also shown in Figure 5. Propylene oxide is the
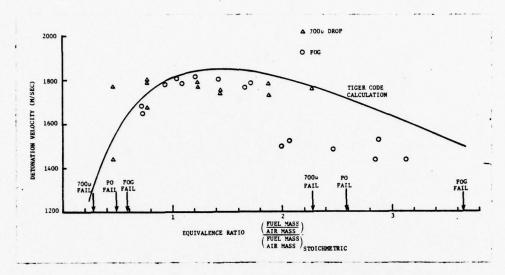
Figure 5. Comparison of Experimental and Theoretical Detonation
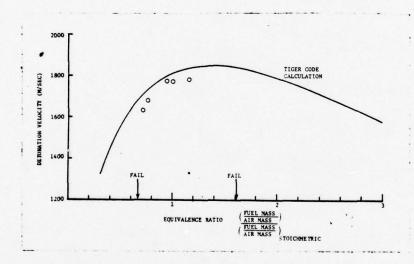Velocities for Various Heptane + 20% NPN/Air Mixtures.

Figure 6. Comparison of Experimental and Theoretical Detonation
Velocities for Various Heptane Fog/Air Mixtures.

standard fuel used in all current fuel – air explosive munitions.
Note that the limits of propylene oxide are narrower than that of
heptane + 20% NPN mixtures. The heptane + 20% NPN fuel should be
able to perform just as well or even better than the propylene oxide
in FAE munitions. This has been confirmed by field tests.

10

Detonations were obtained in an average 23 ft. diameter, unconfined heptane + 20% NPN cloud in air that was formed by explosive dissemination. The average wave speed was 1.58 mm/μsec. For a stoichiometric mixture, on an equal weight basis, the heptane + 20% NPN requires 27.7% more air than propylene oxide. Thus the heptane fuel has the potential of larger area coverage in FAE munitions.

Spark schlieren photographs of the detonation front in the heptane-NPN-air mixture are shown in Figure 7 to illustrate the structure of the wave front. In general, the wave is warped and apparently thick in a schlieren sense. The degree of non-planarity and the thickness of the front progressively increase as the drop size increases. Evidently, this results from poorer coupling between the shock-front and the combustion region in the long reaction zone produced by large drops. The wave front of the 1400 μm drop mixture is slanted so severely that it is actually propagating in a spinning mode as confirmed by periodically oscillating pressure traces produced by this detonation. Transverse waves are visible in all of these pictures. The number of transverse waves decreases as the drop size increases. An anology to the behavior of gaseous detonations near their detonation limits is evident.
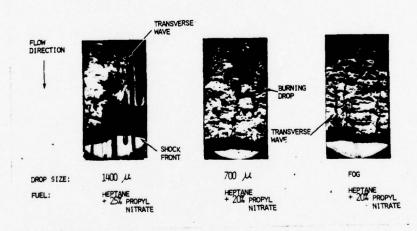

Figure 7. Schlieren Photographs of Two-Phase Detonation Fronts.

Further analogies between heterogeneous phase and gas phase are indicated. As the drop size decreases from 1400 μm to 10 μm, blast waves, generated by localized gaseous volumes behind the incident shock become more evident. Similar kinds of blast waves are observed in gaseous detonations having long reaction zones (Ref 10). These local explosions do not necessarily occur near fuel drops and are caused by fuel vaporized from falling fuel drops prior to detonation initiation. Thus the effect becomes more noticeable as the drop size decreases since greater vaporization occurs. It is interesting to note that, in spite of all indications of gas phase reactions, it has been experimentally verified that the fuel vapor alone is not sufficient to sustain a detonation in the tube.

Another interesting feature is, in contrast to studies conducted in oxygen-rich atmospheres where local blast waves from the burning micromist in the wake of individual fuel drops are observed, (Ref 11), no local blast waves originate from the wake of individual drops in air. This observation agrees with results from the single drop studies. The role of these blast waves in the two-phase detonation reaction zone has been discussed by many investigators (Ref 12). Apparently the presence of local blast waves from the fuel drops depends on the oxygen concentration and the occurrence of these blast waves is not a necessary condition for propagation of two-phase detonations.

CONCLUSIONS

Systematic studies of the effect of additives and fuel drop size on the detonability of heptane-air mixtures have been carried out under controlled laboratory conditions and in large scale field tests. It was shown in the shock tube studies that n-propyl nitrate and butyl nitrite as well as, to a lesser extent, small drop size can greatly widen the detonation limits and reduce initiation requirements of heptane-air mixtures. Large scale field tests of explosively disseminated fuel-air clouds confirm findings obtained in the laboratory tests, demonstrating that systematic laboratory tests can be used to predict detonability and performance of any new fuel system to be used in FAE munitions.

Unique schlieren photographs of the reaction zone of propagating fuel-air detonations have been successfully obtained. No blast waves are observed either in the wake of single shocked fuel drops or from drops in the reaction zones of propagating detonations

with air. Apparently, the assumption which has appeared in the liter-
ature, that such blast waves are necessary to maintain a two-phase
detonation, is incorrect. Furthermore, the insight into the structure
of the reaction zone that these photographs have now made available can
be used to suggest new fuel additives which will affect detonability,
either increasing or decreasing it.

Sensitized heptane is potentially superior to fuels used in
current FAE munitions, providing increased explosive performance
(larger area coverage per unit weight of fuel), greater safety in trans-
port and handling and lower procurement costs. These tests have estab-
lished this system as a desirable alternative to fuels currently being
used.

ACKNOWLEDGEMENTS

13

REFERENCES

1. Lu, P. L. and Slagg, N. (1972), "Chemical Aspects in Shock Initiation of Fuel Droplets", Astronautica Acta, 17, 693-702.

2. Lu, P. L. and Slagg, N. (1974), "Chemical Aspects in Shock Ignition of Fuel Drops", Acta Astronautica 1, 1219-1226.

3. Kauffman, C. W. and Nicholls, J. A. (1971), "Shock Wave Ignition of Liquid Fuel Drops", AIAA J. 9, 880-885.

4. Fishburn, B. (1974), "Boundary Layer Stripping of Liquid Drops Fragmented by Taylor Instability", Acta Astronautica 1, 1267-1284.

5. Calvert, J. G. and Pitts, J. N., Jr. (1966), Photochemistry, 824-829, John Wiley and Sons, N.Y.

6. Dabora, E. K. (1975), "Effect of NO_2 on the Ignition Delay of CH_4-Air Mixtures", Combustion and Flame, 24, 181-184.

7. Kauffman, C. W. , Nicholls, J. A., and Olzmann, K. A. (1971), "The Interaction of an Incident Shock Wave with Liquid Fuel Drops", Combustion Science and Technology, 3, 165-178.

8. Williams, F. A. (1976), "Quenching Thickness for Detonations", Combustion and Flame, 26, 403-406.

9. Vanta, E. B., Parson, G. H. and Collins, P. M.., (1973), "Detonability of Propylene Oxide/Air and n-Propyl Nitrate/Air Mixtures", Technical Report AFATL-TR-73-3, Air Force Armament Laboratory, Eglin Air Force Base, Florida.

10. Lu, P. L., Dabora, E. K., and Nicholls, J. A. (1969), "The Structure of H_2-CO-O_2 Detonations", Combustion Science and Technology, 1, 65-74.

11. Dabora, E. K., Ragland, K. W. and Nicholls, J. A., (1969), "Drop Size Effects in Spray Detonations", Twelfth Symposium (International) on Combustion, 19-26.

12. Borisov, A. A., Gelfand, B. E., Gubin, S. A., Kogarko, S. M., and Podgrebenkov, A. L., (1970), "The Reaction Zone of Two-Phase Detonations", Astronautica Acta 16, 411-417.

A NONLINEAR CONSTITUTIVE RELATIONSHIP
FOR COMPOSITE PROPELLANTS

DONALD L. MARTIN, JR., PhD.
PROPULSION DIRECTORATE, TECHNOLOGY LABORATORY
US ARMY MISSILE RESEARCH AND DEVELOPMENT COMMAND
REDSTONE ARSENAL, ALABAMA 35809

INTRODUCTION

The mechanical response and the failure of composite solid
propellants are known to be related to the formation and growth of
vacuoles on the microscopic or macroscopic scale. Failures in com-
posite materials such as composite propellants originate at the filler
particle or binder molecular level. These microscopic failures,
whether adhesive failures between the binder and filler particles or
cohesive failures in the binder, result in vacuole formation. The
cumulative effect on all vacuole formation and growth is observed on
the macroscopic scale as strain dilatation. The strain dilatation in
composite propellants varies with binder type, binder formulation,
filler particles, and presumably filler particle size distribution.
The strain dilatation caused by vacuole growth and formation results
in a nonlinear stress-strain behavior for these materials.

The experimental investigations of Leeming et al. (1),
Bornstein (2), and Martin (3) emphasize the fact that the state-of-the-
art linear viscoelastic theory does not predict propellant grain
stresses and strains under realistic loading conditions. The most
significant portion of the error is due to the constitutive nonlinear-
ity of the material. Linear thermorheological simple viscoelastic
theory has certain important limitations; however, it is still the only
constitutive theory currently used in most analyses because no other
constitutive equations previously existed that are as easily incorpo-
rated in the analyses.

There has been considerable progress in the development of nonlinear constitutive theories and methods of approximating specific nonlinear viscoelastic behavior in the last few years. There will be no attempt in this paper to review the extensive literature on the general theories of nonlinear viscoelasticity.

BACKGROUND AND THEORY

Farris (4) demonstrated that the mechanical response and the failure in solid propellants are related to the formation and growth of vacuoles within the composite material and that this phenomenon is directly related to the macroscopically observed strain dilatation. The mechanical properties, particularly Poisson's ratio, are strongly dependent on the dilatational behavior of these materials. Strain dilatation measurements, even for the uniaxial tensile tests, are cumbersome, tedious, and time consuming. Therefore, a relationship that would permit the determination of strain dilatation and Poisson's ratio as a function of strain utilizing only the stress-strain data would be most desirable. Insight into the factors which govern the nonlinear response and failure of composite propellants can be obtained utilizing a model for dewetting, vacuole formation, and growth of vacuoles which is based on assumptions regarding the microstructural behavior of the system. The model can then be used to develop an expression for dilatation as a function of stress and strain and to compare the predicted dilatation with experimental measurements of dilatation. The degree of agreement between the predicted and experimental values of dilatation gives an indication of the validity of the model. This approach was used by Farris (5,6), Fedors and Landel (7), and others (8,9).

The present state of the art of predicting dilatation as a function of strain is limited because there is very little quantitative information available on the experimentally determined dilatation under conditions similar to those encountered with solid propellant rocket motors. Various models describing this phenomenon have appeared in the literature. In developing such a model, attention is focused on a rigid filler particle contained in an elastomeric matrix. As the material is strained above some critical value, a vacuole is formed about the filler particle due to the internal failure of the composite. The source of vacuole formation may be either adhesive failure of the filler-binder bond or cohesive failure in the binder near the filler particle (10,11). As the strain in the material increases, the vacuoles continue to increase in size and number. The shape and instantaneous behavior of these vacuoles seem to follow these assumptions:

1) Vacuoles from arbitrarily at any magnitude of strain above the critical value.

2) Each vacuole behaves as an ellipsoid of revolution with the minor axis determined by the diameter of the enclosed filler particle.

3) The major axis of the ellipsoidal vacuole increases linearly with strain at a rate proportional to the size of the filler particle it contains.

Various constitutive theories exist and experimental investigations can be designed to produce specific stress-strain equations. In this section a constitutive relationship for uniaxial tensile data is discussed and the similarity with Farris' equation is noted (4).

Martin (12) used the constitutive relationship

$$\sigma(\epsilon, T, t) = \left\{ F(\epsilon, T, t) e^{-\ell n f(\epsilon)} \right\} \epsilon \quad . \tag{1}$$

Equation (1) has been verified experimentally.

Starting with Farris' equation and defining $F = \sigma/\epsilon$ as the secant modulus, as was used by Martin, the following relationship is obtained:

$$F(\epsilon) = \frac{\sigma}{\epsilon} = E_1 \, e^{-\beta \left(\frac{\Delta V}{V_o} \right) 1/\epsilon} \quad , \tag{2}$$

where E_1 is the initial modulus and $F(\epsilon)$ is meant to indicate F at a given strain. This value will be indicated simply by F in the remainder of this work. After rearrangement one obtains:

$$\frac{F}{E_1} = e^{-\beta(\Delta V/V)1/\epsilon} \quad . \tag{3}$$

Equation (3) indicates that the nonlinearity of the stress-strain curve (of which the vacuole formation and growth is considered to be the major contributor) should be described by the ratio of F/E_1. Taking the natural logarithm of both sides of Equation (3) results in Equation (4):

$$\beta \left(\frac{\Delta V}{V} \right) \frac{1}{\epsilon} = \ell n \left(\frac{E_1}{F} \right) \quad . \tag{4}$$

Since β is assumed to be independent of strain magnitude, Equation (3) suggests that a plot of $\{\epsilon\ln(E_1/F)\}$ versus strain should have a similar shape to the dilatation-strain curve. Figure 1 shows plots of $\{\epsilon\ln(E_1/F)\}$ and $(\Delta V/V)$ versus strain. The similarity in the shapes of
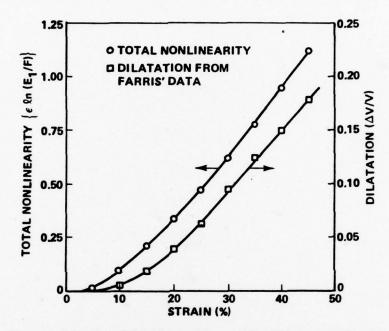

Figure 1. Total Nonlinearity and Dilatation Versus Strain for a Granular Filled Elastomer.

the two curves indicates that the total nonlinearity, $\{\epsilon\ln(E_1/F)\}$, in the materials' behavior may be calculated from an equation similar to that used by Farris (4). The same reasoning used by Farris results in the expression

$$f(\epsilon) = S\gamma_{max} \int_{-\infty}^{n} \int_{-\infty}^{n} \frac{e^{-n2/2}}{\sqrt{2\Pi}} \, dndn + \gamma_0\epsilon \quad , \quad (5)$$

where

$f(\epsilon)$ = total nonlinearity of the materials' behavior

$S = (A - B)\sqrt{2\Pi}/\gamma_{max}$

γ_o = initial slope of the $\{\epsilon \ell n(E_1/F)\}$ versus strain curve

γ_{max} = final slope of the $\{\epsilon \ell n(E_1/F)\}$ versus strain curve

$\bar{\epsilon}$ = the strain magnitude at the intersection of the initial asymptote of slope γ_o and the final asymptote of slope γ_{max}

A = the magnitude of $\{\epsilon \ell n(E_1/F)\}$ at $\bar{\epsilon}$

B = the magnitude of the ordinate at the intersection of the two asymptotes

$n = (\epsilon - \bar{\epsilon})/S$.

Equation (5) then would permit one to estimate the total nonlinearity of the propellant once the parameters have been determined. It would be more useful, however, to obtain a quantity representing the propellant's nonlinear behavior that could be readily utilized in the current linear viscoelastic programs to account for the nonlinear response of the material. With this in mind, consider Equation (3), which after rearranging becomes

$$\beta\left(\frac{\Delta V}{V}\right) = \epsilon \ell n\left(\frac{E_1}{F}\right) \quad . \tag{6}$$

Recalling that β and E_1 are independent of strain and differentiating both sides with respect to strain, one has

$$\beta \frac{d\left(\frac{\Delta V}{V_o}\right)}{d\epsilon} = \ell n \frac{E_1}{F} + \left(1 - \frac{E}{F}\right) \quad , \tag{7}$$

where E = the instantaneous slope of the stress-strain curve. Equation (7) now contains two unknowns, β and $d/d\epsilon(\Delta V/V_o)$. Now, consider the change of dilatation with respect to strain. According to Farris' data (5), the quantity $d/d\epsilon(\Delta V/V_o)$ obtains a maximum somewhere between the points of maximum stress and rupture stress and remains constant until failure occurs. Farris proposed the relationship

$$\frac{d}{d\epsilon}\left(\frac{\Delta V}{V_o}\right) = c \, V_{fd} \quad , \tag{8}$$

where V_{fd} = volume fraction of dewetted solids. Experimental data were presented for systems that tend to dewet completely. These data gave excellent agreement with the model represented by Equation (8). The maximum slope of the dilatation-strain curve is shown to be directly proportional to the volume fraction of filler particles in the material. For materials that tend to dewet completely, the proportionality factor is unity (c = 1); therefore,

$$\left[\frac{d}{d\epsilon} \left(\frac{\Delta V}{V_o} \right) \right]_{max} = \phi_f \quad , \tag{9}$$

where ϕ_f is the total volume fraction of filler particles in the system.

With these observations noted, the following simplifying assumptions will be made that will enable the determination of the factor β without having to conduct the dilatation-strain measurements:

1) The maximum slope of the dilatation-strain curve is obtained at a strain magnitude equal to halfway between the point of maximum stress and breaking stress.

2) All composite propellants are totally dewetted in the region of the sample where failure occurs.

The assumption that all filler particles in composite propellants are totally dewetted in the region of failure is not too unrealistic for the following reasons. In composite propellants where the bond between the filler particles and the binder is stronger than the cohesive strength of the binder, vacuoles initiate in the binder and propagate to the surface of the filler particle prior to failure. In these materials there is probably a large strain gradient in the region of failure and there would be a high probability of totally dewetted filler particles in this region. One reason that Equation (9) does not appear to be true for all propellants is that the measured dilatation, whether localized or not, is averaged over the total volume of the sample. If the localized band of dewetted filler particles is narrow, the measured dilatation would appear to be much smaller than the actual dilatation in the region of failure. The maximum slope of the dilatation-strain curve is related to the dewetted filler particles, but the dewetted filler particles may be contained in a narrow band around the failure region.

MARTIN

The assumption that the maximum slope of the dilatation-strain curve is obtained at a strain magnitude halfway between the points of maximum stress and breaking stress is approximately correct according to Farris' data. The purpose of this assumption is to fix the point on the stress-strain curve for the calculation of the coefficient β. One reason that failure points are not recommended for this calculation is the variation in ways of selecting failure points. The use of the stress and strain values at $\epsilon = (\epsilon_m + \epsilon_b)/2$ should minimize the variations in methods of selecting failure points as a source of error in the calculation of β.

Substituting Equation (9) into Equation (7) yields the following results after rearranging:

$$\beta = \frac{1}{\Phi_f}\left\{ \ell n \, \frac{E_1}{F} + 1 - \frac{E}{F} \right\} \quad , \tag{10}$$

to be evaluated at $\epsilon = (\epsilon_m + \epsilon_b)/2$.

With β known, one can then utilize Equation (6) to determine the dilatation as a function of stress and strain. The resulting expression is as follows:

$$\left(\frac{\Delta V}{V_o}\right) = \frac{\epsilon}{\beta} \, \ell n \, \frac{E_1}{F} \quad . \tag{11}$$

This equation was utilized to calculate the apparent volume change for two composite materials where the experimental dilatation-strain data were available. The calculated and experimentally measured dilatation-strain data were in excellent agreement.

To more readily utilize the volume change data obtained in this way for a nonlinear viscoelastic stress analysis, the following method is proposed. It would be more desirable to reflect the propellant nonlinearity in an equivalent Poisson's ratio, which would be allowed to vary with strain magnitude, strain rate, and temperature. The variable Poisson's ratio could then be incorporated into the current linear viscoelastic stress analysis program to reflect the propellant's nonlinear behavior.

To derive an expression of Poisson's ratio as a function of stress and strain, the volume change is approximated by the first strain invariant neglecting higher order terms involving strain.

$$\left(\frac{\Delta V}{V_o}\right) = e_1 + e_2 + e_3 \quad , \tag{12}$$

where e_1, e_2, and e_3 are the strains in the principal directions and are related as follows in the uniaxial condition:

$$e_1 = \epsilon$$

$$e_2 = e_3 = -\nu e_1 = -\nu\epsilon \quad , \tag{13}$$

where ν is Poisson's ratio. Substituting these relationships into Equation (11) results in the following

$$\left(\frac{\Delta V}{V_o}\right)\frac{1}{\epsilon} = (1 - 2\nu) \quad . \tag{14}$$

Substituting Equation (14) into Equation (6) and rearranging, one obtains the relationship:

$$\nu = \frac{1}{2}\left\{1 - \frac{1}{\beta} \ln\left(\frac{E_1}{F}\right)\right\} \quad . \tag{15}$$

Also, substituting Equation (14) into Farris' equation yields the following constitutive relationship for the uniaxial nonlinear behavior:

$$\sigma = E_1\epsilon e^{-\beta(1-2\nu)} \quad . \tag{16}$$

For the nonlinear viscoelastic treatment, the initial slope modulus, E_1, is allowed to vary with strain rate and temperature; the non-linearity coefficient, β, is allowed to vary with strain rate and temperature; and the Poisson's ratio, ν, is allowed to vary with strain rate, temperature, and strain magnitude.

EXPERIMENTAL VERIFICATION

The model used in deriving the expressions for dilatation and Poisson's ratio as a function of stress and strain was based on certain assumptions regarding the microstructural behavior of composite materials such as composite propellants. The validity of the assumptions and the model is indicated by good agreement of calculated and experimental stress and dilatation data.

Figures 2, 3, and 4 are plots of stress and dilatation versus strain for a granular filled elastomer. The experimental data were obtained from Farris'work (5). The experimental data in these figures were obtained at 25°C (75°F) and at strain rates of 6.66, 66.6, and 666%/min. The plotted points represent the data read from Figure 6 in Farris' paper. The solid lines indicate the values calculated using Equation (16) for stress and Equation (11) for dilatation. Figures 2, 3, and 4 indicate good agreement between calculated and experimental values of stress and dilatation at all strain rates. The dilatation-strain data appear to be shifted by a constant factor along the strain axis in some cases. This would indicate a slight error in the dilatation measurements, which could be attributed to localized dewetting or vacuole formation as discussed previously. Equation (16) is shown to yield excellent agreement with experimental stress values at all strain rates. In Figure 4, Equation (11) is shown to also give excellent agreement with the dilatation experimental data at a strain rate of 666%/min up to a strain magnitude of 35%. At strain magnitudes in excess of 35%, the experimental dilatation data fall below the predicted values. While vacuoles may form arbitrarily at all strain magnitudes above a critical value, it is reasonable to assume that the weakest point in the composite material will be more susceptible to vacuole formations and should contain a larger population of vacuoles than the remainder of the material. The formation of vacuoles to relieve the large strain gradient that would be present at the weakest point in the composite requires a finite time in which to react. If the composite material is strained at a rate faster than the material can respond through microstructural failure and vacuole formation, the dilatation would appear to be uniform throughout the material until conditions were such that the rate of dilatation at the weakest point in the material became large enough to promote the formation of a localized band of a large population of vacuoles. The data presented in Figure 4 indicate that at a strain rate of 666%/min, near uniform dilatation is present up to a strain magnitude of approximately 35%. At this strain a localized band of a large population of vacuoles developed at the weakest point in the composite material, which then began to have a strong influence on the material's behavior. As the material is strained in excess of 35%, the volume increase in the vacuoles within the established localized failure region is more than the measured average dilatation. This would explain why the experimental dilatation data presented in Figure 4 became less than the calculated values as the strain magnitude increased to exceed 35%. At the lower strain rates of 6.66 and 66.6%/min, the localized failure region was established in the initial portion of the test ($\epsilon < 0.10$), thus explaining why experimental dilatation deviated slightly from the calculated dilatation for the duration of the test.
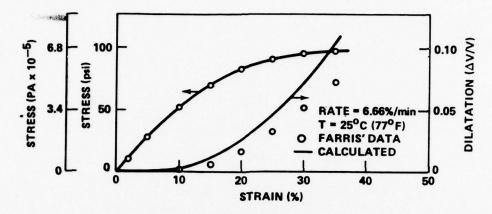
Figure 2. Dilatation and Stress Versus Strain
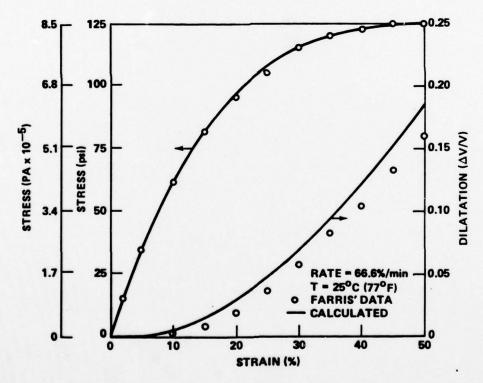for a Granular Filled Elastomer.

Figure 3. Dilatation and Stress Versus Strain
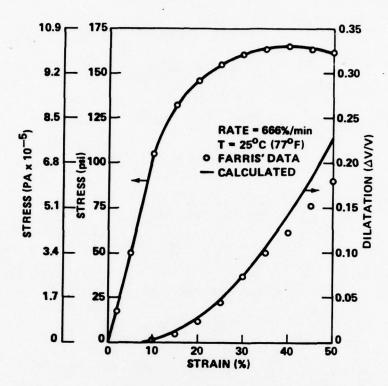for a Granular Filled Elastomer.

Figure 4. Dilatation and Stress Versus Strain
for a Granular Filled Elastomer.

Figures 2, 5, and 6 are plots of stress and dilatation
versus strain. The experimental data in these figures were obtained
from Figure 3 in Farris' paper (5). These data indicate the stress
and dilatation behavior of a typical granular filled elastomer at a
strain rate of 6.66%/min at 2 5°, 4.4°, -18°C (77°, 40°, and 0°F).
The data shown on Figures 2, 5, and 6 indicate excellent agreement
between values of stress calculated by Equation (15) and the experi-
mental values at all temperatures. The behavior of the dilatation
strain data is similar to that shown on Figures 2, 3, and 4 with the
experimental values deviating slightly from the values calculated by
Equation (11). The same reasoning as used previously would also
explain the deviation of the experimental dilatation from that cal-
culated by Equation (11), as indicated on Figures 2, 5, and 6.

Normally, nominal stress values were corrected to true
stress by multiplying the nominal stress value by the extension
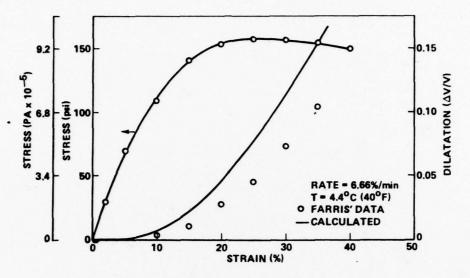ratio, λ_1. This correction was originally derived based on the

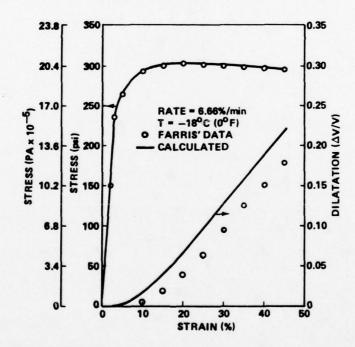Figure 5.   Dilatation and Stress Versus Strain
for a Granular Filled Elastomer.

Figure 6.   Dilatation and Stress Versus Strain
for a Granular Filled Elastomer.

MARTIN

assumption that there is no volume change within the sample for the duration of the test. The relationship considering volume change is derived as follows:

$$\left(\frac{\Delta V}{V_o}\right) = \lambda_1 \lambda_2 \lambda_3 - 1 \quad , \tag{17}$$

where $\lambda_i = 1 + \epsilon_i$. Therefore, rearranging Equation (17), one obtains

$$\lambda_2 \lambda_3 = \frac{1 + \left(\frac{\Delta V}{V_o}\right)}{\lambda_1} = \frac{A}{A_o} \quad . \tag{18}$$

Substituting this expression into the equation $\sigma_c = \sigma_o A_o/A$, one obtains the following:

$$\sigma_c = \frac{\sigma_o \lambda_1}{1 + \frac{\Delta V}{V_o}} \quad . \tag{19}$$

Therefore, the true stress σ_c is lower than would be obtained by the normal correction procedures. Equation (19) reduces to the equation normally used as ΔV approaches zero.

CONCLUSIONS

A nonlinear constitutive relationship is developed for uniaxial stress-strain conditions that are shown to fit experimental data over a wide range of strain rates and temperatures. The total stress-strain nonlinearity of particulate filled composites may be represented by statistical parameters similar to those used by Farris to describe the dilatation-strain behavior. By using Equations (10) and (11), the dilatational behavior of composite propellants may be predicted from the stress-strain data and the total filler content. The relationships derived in this report will enable one to investigate the nonlinear behavior at strain-rates and temperatures where it is very difficult, if not impossible, to obtain measured dilatation data.

The relationship represented by Equation (15) may be utilized, along with experimental stress-strain data obtained at different strain-rates and temperatures, to determine the dependency of an equivalent Poisson's ratio on the strain-rate, temperature, and strain magnitude. The Poisson's ratio thus determined may then be utilized

in currently available computer programs to describe transient thermo-viscoelastic behavior with the consideration of the materials' non-linear response.

The relation normally used for obtaining corrected stress data by multiplying the nominal values by the extension ratio would yield values of stress and strain at the maximum corrected stress, which were significantly larger than the true stress-strain values would indicate. These errors could conceivably result in the wrong conclusions as to the structural integrity of marginal designs; i.e., one might predict a safe design utilizing stress-strain allowables based on the normal correction method when, in reality, the true stress-strain allowables would indicate the design would be expected to fail. It is therefore recommended that stress-strain allowables for use in future failure analysis be based on true stress as defined by Equation (18).

REFERENCES

1. Leeming, H., et al., <u>Solid Propellant Structural Test Vehicle and System Analysis</u>, LPC Final Report No. 966-F (AFRPL-TR-70-10), Lockheed Propulsion Company, Redlands, California, March 1970.

2. Bornstein, G. A., "Transient Thermoviscoelastic Analysis of a Uniaxial Bar," <u>Bulletin of 8th Mechanical Behavior Working Group</u>, Vol. 1, CPIA Publication No. 177, 1968.

3. Martin, D. L., Jr., "An Approximate Method of Analysis of Non-linear Transient Thermoviscoelastic Behavior," <u>8th Meeting of JANNAF Mechanical Behavior Working Group</u>, Vol. 1, CPIA Publication No. 193, 1970.

4. Farris, R. J., "The Influence of Vacuole Formation on the Response and Failure of Solid Propellants," <u>Trans. Soc. Rheol.</u>, Vol. 12, 1968.

5. Farris, R. J., "Strain Dilatation in Solid Propellants," <u>Bulletin of 3rd Meeting of ICRPG Mechanical Behavior Working Group</u>, CPIA Publication No. 61U, Vol. 1, October 1964, p. 291.

6. Farris, R. J., "Dilatation of Granular Filled Elastomers Under High Rates of Strain," <u>J. of Applied Polymer Science</u>, Vol. 8, 1964, pp. 25-35.

END
DATE
FILMED

8 -78

DDC

7. Fedors, R. F. and Landel, R. F., "Mechanical Behavior of SBR-Glass Bead Composites," JPL Space Program Summary 37-41, Vol. IV.

8. Elliott, D. M. and Ledbetter, W. B., "Microstructural Damage and Its Influence on the Dilatation-Strain Properties of Filled Elastomeric Composite Materials," Bulletin of the 6th ICRPG Mechanical Behavior Working Group, CPIA Publication No. 158, Vol. 1, 1967, pp. 235-251.

9. Jones, H. C. and Yiengst, H. A., "Dilatometer Studies of Pigment-Rubber Systems," J. of Industrial Engineering Chemistry, Vol. 32, No. 10, 1940.

10. Oberth, A. E. and Bruenner, R. S., "Binder-Filler Interaction and Propellant Physical Properties," Bulletin of 4th ICRPG Mechanical Behavior Working Group, October 1965, p. 45.

11. Hilzinger, J. E., "Microstructural Behavior of Filler Reinforced Elastomeric Binders," Bulletin of 5th ICRPG Mechanical Behavior Working Group, CPIA Publication No. 119, Vol. 1, 1966, p. 169.

12. Martin, D. L., Jr., "An Approximate Method of Analysis of Nonlinear Transient Thermoviscoelastic Behavior," PhD Dissertation, Department of Mechanical Systems Engineering, University of Alabama, University, Alabama, 1968.

RADAR ABSORPTIVE MATERIAL FORM INDUSTRIAL EFFLUENT

*JAMES W. McCAULEY, PhD
BERNARD M. HALPIN, JR., PhD
ARMY MATERIALS AND MECHANICS RESEARCH CENTER, WATERTOWN, MA   02172
STEPHEN D. EITELMAN, MR.
U.S. ARMY MATERIEL DEVELOPMENT AND READINESS COMMAND
ALEXANDRIA, VA   22333
THOMAS V. HYNES, PhD
ARMY MATERIALS AND MECHANICS RESEARCH CENTER, WATERTOWN, MA   02172

I.   INTRODUCTION

Recent innovative work (1,2) by the Nippon Electric Co. (NEC) Tokyo, Japan, on a ferrite precipitation method for removing heavy metal ions from industrial waste water and the subsequent potential use of this material (ferrite sludge) as a cheap microwave absorber has stimulated interest in the possible United States military applications of this concept. An important factor in this consideration is that the cost of the material is projected as significantly less than currently available ferrite material. This report describes the results of this program: a characterization and fabrication phase carried out at the Army Materials and Mechanics Research Center, Watertown, Massachusetts, and a testing phase carried out at the U.S. Army Electronic Proving Ground, Fort Huachuca, Arizona.

II.   MAGNETIC DIELECTRIC RADAR SIGNATURE REDUCTION

One way of reducing radar signatures of highly reflective materials is by the use of radar-absorbing coating materials. Magnetic dielectrics (ferrites) are a family of materials that can absorb microwaves by the lossy interaction of the electric and magnetic vectors of an electromagnetic wave with the material.

*McCAULEY, HALPIN, EITELMAN
 and HYNES

When an electromagnetic wave strikes an object there is generally a reflection because of a discontinuity in electrical properties between the medium the wave is travelling and the object it strikes. One can show the following: an electromagnetic wave of amplitude E_0 travelling in free space (μ [the magnetic permeability] $= \mu_0$, and ε [the electrical permittivity] $= \varepsilon_0$) and striking a plane surface A ($\mu_A = \mu_A' + i\mu_A''$ and $\varepsilon_A = \varepsilon_A' + i\varepsilon_A''$, where $\mu_A' \gg \mu_A''$ and $\varepsilon_A' \gg \varepsilon_A''$) will have a relection approximately $[1-[(\mu_A'/\varepsilon_A')(\varepsilon_0/\mu_0)]^{\frac{1}{2}}$ in magnitude. If $\mu_A'/\varepsilon_A' = \mu_0/\varepsilon_0$, there will be no reflection. The wave will be attenuated in the material by a factor $\exp(-\mu_A'' \omega[\mu_A'/\varepsilon_A']^{-\frac{1}{2}}\ell)$ where $\omega = 2\pi$ frequency and $\ell =$ the depth in the material. Thus by properly choosing the permeability, dielectric constant, depth, and loss factors, normal incident waves can be absorbed in the material. Unfortunately, however, all factors vary with frequency and this factor must be taken into consideration in the exact calculations. In this investigation, microwave attenuation and concomitant radar signature reductions have been achieved by an increase in the magnetic permeability and magnetic losses of a dielectric material.

III. DESCRIPTION OF FERRITE PRECIPITATION PROCESS

Stimulated by more stringent pollution control requirements in Japan, the NEC is investigating the use of a ferrite precipitation process to remove heavy metals from their industrial waste water (1). They have two motives for exploring this unique idea: (1) purification of industrial waste water and (2) simultaneous formation of significantly cheaper ferrite raw material for possible commercial use. Potential uses envisioned by them include prevention of double TV images (ghosts) by using ferrite coatings on buildings and also elimination of ground reflections near radar receivers by coating the ground.

Standard industrial practice (3) used in the production of ferrite materials consists of elaborate mixing, pressing, and controlled heat treatments of carefully weighed, mechanically pulverized, and mixed ceramic oxide powders. The novel technique being used by NEC is a wet chemical reaction process referred to as coprecipitation. Essentially, divalent iron and other metal ions are hydrolyzed in a basic solution and then oxidized into the ferrite spinel structure (2).

The two steps can be formularized as follows:

Hydrolysis:

$$M_{1-x}^{2+} + Fe_{2+x}^{2+} + 6OH^- \rightarrow M_{1-x}Fe_{2+x}(OH)_6 \qquad (1)$$

*McCAULEY, HALPIN, EITELMAN
and HYNES

Warm Aeration:

$$M_{1-x}Fe_{2+x}(OH)_6 + \frac{1}{2} O_2 \rightarrow M_{1-x}Fe_{2+x}O_4 + 3H_2O \qquad (2)$$

or more simply as

$$Fe^{2+} + M^{2+} + ROH + O_2 \rightarrow M_{1-x}Fe_{2+x}O_4 \qquad (3)$$

where R = alkali or alkali earth cation (ex. NaOH).

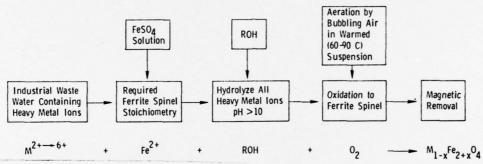In summary, the ferrite sludge coprecipitation process is schematically illustrated in Figure 1.

Figure 1. Schematic flow chart for the production and removal of ferrite sludge by the coprecipitation process.

Experimental data demonstrate that the heavy metal removal technique is successful. Table 1 lists various heavy metal ions in the industrial waste water before and after the coprecipitation treatment (1).

IV. CHARACTERIZATION OF FERRITE SLUDGE

Two separate batches of sludge, nominally referred to as the 5-kg and 50-kg batches, were obtained from NEC. They were brown to dark brown in color and had a mud-like consistency. Figure 2 illustrates both the as-received material and also the dried, ball-milled, and sieved (-325 mesh) powder.

Table 1. HEAVY METAL ION CONCENTRATIONS IN INDUSTRIAL WASTE WATER BEFORE AND AFTER FERRITE SLUDGE COPRECIPITATION TREATMENT.

Heavy Metal Ion	Concentration in Waste Water (ppm)	
	Before Treatment	After Treatment
Cu	9,500	< 0.5
Ni	20,000	< 0.5
Sn	4,000	<10
Pb	6,800	< 0.1
Cr^{6+}	2,000	< 0.1
Cd	1,800	< 0.1
Fe	over 1%	< 1.0
Hg	3,000	ND*

*Nondetectable

Figure 2. Ferrite sludge.

AS RECEIVED

DRIED, BALL MILLED,
AND SIEVED

The as-received sludge contains appreciable amounts of water and must be thoroughly dried prior to further use. Table 2 lists the water content and approximate chemistry of the two batches which are quite different in composition, the 5-kg batch being iron-rich and the 50-kg batch being iron-poor.

Hypothetical chemical-structural formulas for these ferrite spinels are as follows (4,5):

5-kg batch: $M'_x Fe^{2+}_{1-x} |M''^{IV}_y Fe^{3+}_{2-y}|^{VI} O_4$

50-kg batch: $M' |M''^{IV}_x Fe^{3+}_{2-x}|^{VI} O_4$

M', M'' = heavy metal ions,

IV, VI = tetrahedral and octahedral

coordination, respectively.

Table 2. CHEMISTRY OF AS-RECEIVED FERRITE SLUDGE

Element	Weight Percent in Batches	
	5 kg	50 kg
H_2O*	45.0	43.9
Fe†	65.7	49.2
Mn‡	0.5-1	1
Mg	.5	1
Al	.1-0.5	1
Ni	.5	1
Si	.5	1-10
Pb	.1-0.5	0.5-1
Sn	.05	.5
Zn	.1	.5-1
Cr	.01	.5
Ca	.1	.5
Cd	.5	.5
Cu	.05	.05-0.1
Mo	.05-0.1	.05-0.1
Ti	.01-0.05	.1-0.5
Zr	.05	.05
Ba	.05	.05

*Loss on drying at 105 C.
†Quantitative analysis for Fe by wet chemistry
‡All the remaining analysis are semiquantitative emission spectroscopic results reported in weight percent

There are eight formula units in the ferrite spinel unit cell, yielding the following actual chemical formula: $M'_8 M''_{8x} Fe_{16-8x} O_{32}$.

Hence, there are at least $M'_8 + M''_{8x}$ crystallographic sites in the unit cell for the removed heavy metal ions.

Figure 3 is a diagrammatic representation of the actual powder diffraction patterns using filtered CrKα radiation. The 5-kg batch is 100% ferrite with a $LiFe_5O_8$ (ferrite)-like structure. The 50-kg sludge is a relatively impure ferrite sample containing several nonferrite phases. Roughly 60% of the mixture is a ferrite similar to $LiFe_5O_8$,

*McCAULEY, HALPIN, EITELMAN
and HYNES

3% is $Fe(OH)_3$, 6% is $ZnCr_2O_4$ (spinel), 11% is NiO, and the last 20%
consists of $FeCl_3$, $CaSnO_4$, and $CaFeO_7$ (ferrite). It is also possible
that a significant percentage of both batches contain X-ray amorphous
phases and therefore do not appear in the diffraction patterns.

Table 3 lists the physical powder characteristics of the two
batches. Figure 4 illustrates a high magnification view of the 5-kg
powder, showing its very fine grain size. The 5-kg batch is a higher
density, larger grain size material than the 50-kg material. Ideal
densities for ferrites of similar chemistry as reported here are about
5.0 g/cc, close to the 5-kg material, but significantly different than
the 50-kg material. All these data indicate that the 5-kg powder more
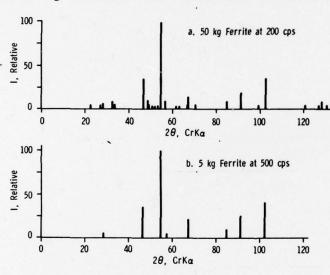
Figure 3. X-ray powder diffraction patterns ($CrK\alpha$) of two dried, ball-milled, and sieved (-325) ferrite sludge powders; $CrK\alpha$ radiation.

Table 3. POWDER CHARACTERISTICS OF FERRITE SLUDGE

Drying Conditions: 150 C, 24 hr

Type kg	Heat Treatment T°C	t(hr)	Atm	Surface Area* m^2/g	Equivalent Spher. Diam.+ μm	ρ_o‡ g/cc
5	-	-	-	20.83	0.058	4.88
5	480	24	Air	16.12	0.075	-
5	480	72	Air	14.26	0.084	-
50	-	-	-	29.97	0.040	3.81
50	400	26	Air	23.77	0.050	-
50	400	136	Air	23.11	0.052	-

*By a BET technique
+Equivalent spherical diameter assuming a density of
 5.0 g/cc
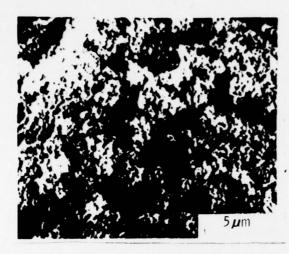‡Measured density by a helium-air pycnometer method

Figure 4. SEM photomicrograph of
fracture surface of ferrite sludge
powder cold isostatically pressed
at 20,000 psi and heat treated
for 24 hr at 200 C.

nearly represents a ferrite material, whereas the 50-kg material is a
very impure ferrite powder, suggesting that the NEC process will not
yield identical material. However, the ferrite sludge in both cases
is a magnetic ferrite, and no heat treatment is required to form a
ferrite phase as in conventional ferrite manufacturing processes.
Further, there seems to be a direct correlation between the iron con-
tent and the ferrite content of the two batches. This implies that
enough $FeSO_4$ must be added to coprecipitate all the heavy metals in a
ferrite phase, otherwise the hydrolysis and warm aeration will produce
other nonferrite material.

V.   FABRICATION PROCEDURE

Three processing routes were analyzed for use in the produc-
tion of the large 12" × 12" testable sample tiles:

a.   pressed powder, no heat treatment,
b.   sintered powder, and
c.   organic matrix composite.

Prior work by the NEC implied that either a or b was suffi-
cient for wave guide (VSWR) microwave absorption measurements on small
samples (6-8). However, because of our remote test location and large
sample requirements for radar absorption measurements only methods b
and c were considered practical. In order to decide between these
procedures a preliminary thermal analysis and sintering study was car-
ried out on the two batches.

a. Thermal Analysis

Initial sintering runs clearly demonstrated a complex sequence of reactions and densification characteristics, the major problem being the oxidation of the ferrite to the nonmagnetic α-Fe_2O_3 (hematite) phase. The reaction can be simplified as follows:

$$2Fe^{2+}Fe_2^{3+}O_4 + \frac{1}{2} O_2 \rightarrow 3Fe_2^{3+}O_3. \tag{4}$$

Figure 5 illustrates the thermogravimetric analysis (TGA) curves for the two batches and Figure 6 shows the differential thermal analysis (DTA) traces of heat-treated and untreated samples. Apparently, because of excessive volatilization of nonferrite metal oxide phases, the TGA curves show weight losses and not weight gains as would be expected for reaction (Eq. 4). However, the DTA curves clearly show the α-Fe_2O_3 reaction at about 604 C in the 5-kg material (Figure 6a) and roughly 340 C in the 50-kg material (Figure 6b). This was confirmed by X-ray diffraction analysis of resultant products. In an attempt to stabilize the ferrite spinel structure, extended heat treatments were carried out below the transformation temperature. The DTA curves for these heat-treated samples (Figures 6c and 6d), however, again show the α-Fe_2O_3 transformation. Clearly, heat treatment in air above the
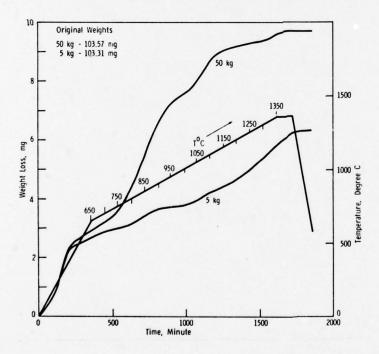

Figure 5. Thermogravimetric analysis traces of predried and pressed ferrite sludge powder.
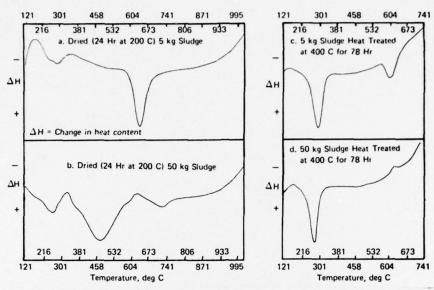
Figure 6. Differential thermal analysis heating curves of
ferrite sludge powders - all heated in air.

α-Fe_2O_3 transformation temperature always results in an α-Fe_2O_3 product
which is unacceptable. One additional DTA experiment using flowing
argon instead of air confirmed this deduction.

b. Sintering Studies

Preliminary sintering work was also carried out with a
summary of the data plotted on Figure 7, showing the variation of
observed densities as a function of temperature. The data simply
reflects the oxidation of the ferrite sludge to α-Fe_2O_3 and its sub-
sequent densification. All samples were isostatically cold pressed
at 20,000 psi prior to heat treatment. Scanning electron microscope
(SEM) photomicrographs of fracture surfaces of runs labeled 18, 23,
and 24 in Figure 7 are presented in Figure 8. As can be seen, sinter-
ing in flowing air results in a completely α-Fe_2O_3 body (Figure 8b),
whereas sintering in argon or a limited amount of air results in a
sintered ceramic containing ferrite (Figure 8a and 8c). Figure 9
shows the microstructural details of the precipitation of α-Fe_2O_3 from
a ferrite matrix. This sample was dried at 200 C for 24 hours and then
heat treated at 1200 C for 24 hours in air. All these data indicate
that the production of large fully-sintered ferrite ceramic tiles must
be preceded by an in-depth atmosphere-controlled sintering study, which
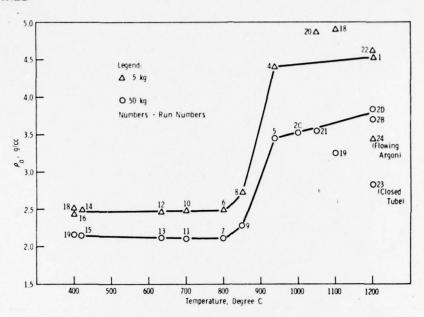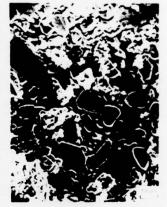was deemed prohibitively expensive as a processing procedure.

Figure 7. Variation of observed density of heat-treated ferrite sludge powder.

a. 100% Ferrite: Calcined in Air at 487 C for 22 Hr, Sintered at 1200 C for 4 Hr in Flowing Argon

(Run 24)

b. 100% α-Fe_2O_3: Calcined in Air at 400 C for 78 Hr, Sintered at 1100 C for 40 Hr in Air

(Run 18)

c. 22% Ferrite, 78% α-Fe_2O_3: Calcined in Air at 400 C for 48 Hr, Sintered at 1200 C for 4 Hr in Closed Muffle Tube

(Run 23)

Figure 8. SEM photomicrographs of representative fracture surface microstructures of heat-treated ferrite sludge powders.

*McCAULEY, HALPIN, EITELMAN
  and HYNES

Figure 9. SEM photomicrograph of fracture surface of ferrite sludge heat treated at 1200 C in air for 24 hours. Microstructure of precipitation of lenticular α-Fe_2O_3 from ferrite matrix.

c. Feasibility of Organic Matrix

It was very quickly demonstrated that at least 50 wt% ferrite powder (-325 mesh) could easily be incorporated into an EPON 828/Z curing agent (100 parts/20 parts) epoxy resin matrix. The powder was not rejected by the resin and formed a homogeneous ferrite/resin composite material and did not lose its magnetic properties.

VI. PRODUCTION OF FERRITE/RESIN TILES

A special mold was constructed to form, de-gas, and cure the ferrite/resin composite. Additionally, it was designed to have a 1/16" overhang of resin over the Al backup plate to prevent direct exposure of the Al plate to the microwave signals, even in a slightly tilted situation. The following is a summary of the final fabrication process.

a. Dry at 150 C and ball mill for at least 24 hours.

b. Sieve all powder through -325 mesh screen.

c. Mechanically blend 50 wt% ferrite powder in an EPON 828/Z curing agent resin mixture.

d. Casting procedures:

(1) Self-bonding

(a) De-gas ferrite/resin mixture.

(b) Cure 2 hr at 175 F, 2 hr at 300 F, directly to Al backup plate.

(2) Two-step process

(a) De-gas ferrite/resin mixture.

(b) Cure 4 hr at 175 F using a mold release on Al backup plate.

(c) Bond to Al backup plate using a two-part Epoxy Adhesive 907.

e. Grind ferrite/epoxy surface to final required thickness dimension.

Four composite plates were fabricated and described in Table 4. Samples FP-1 through FP-4 were all bonded to 12" × 12" × 5/16" Al plates; the ferrite/resin composites measured 12-1/8" × 12-1/8". Side views of these test panels can be seen in Figure 10.

Table 4. DESCRIPTION OF FERRITE/RESIN COMPOSITE PLATES

Sample	Ferrite	Resin Thickness	Process
FP-1	Undoped	1/8"	1 step
FP-2	50 kg-50 Wt%	1/8"	1 step
FP-3	50 kg-50 Wt%	3/8"	2 step
FP-4	5 kg-50 Wt%	1/8"	2 step
FP-5	Blank Al Plate		

Figure 10. Photograph of four test tiles and standard Al plate - side view.

11

VII. MICROWAVE ABSORPTION TESTS

The microwave absorption tests were carried out by Captain
J. Newton, Mr. K. Hakes, Mr. W. Krotzer, and Warrant Officer G.
Van Horn at the Electromagnetic Interference Laboratory, Black Tail
Canyon Test Facility, U.S. Army Electronic Proving Ground, Fort
Huachuca, Arizona. Figure 11 is a drawing of the actual arrangement
of the test panel, receiving antenna, and transmitting horn used in
this experiment. After several unsatisfactory series of measurements
using either a wood or lucite test panel holder, a target holder con-
sisting of a styrofoam block with a recess cut to hold the test panels
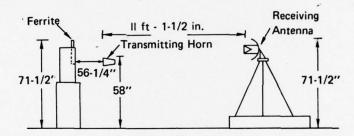was selected for the final measurements.

Figure 11. Physical arrangement of microwave test equipment
and test plate in Fort Huachuca anechoic chamber.

Average ΔdB (difference between the standard uncoated Al plate
and ferrite/resin coated plates) and equivalent percent absorption for
the various test panels are plotted against frequency in Figure 12.
Also plotted on the figure are the common radar bands of interest.
All the microwave measurements were repeatable within 1.5 dB. There
is significant attentuation of the microwave signals by the 50-kg
3/8"-thick ferrite/resin composite in two regions: about 5 dB absorp-
tion (68%) in the 3 to 4.5 GHz region and over 6 dB absorption (75%)
in the 10 to 13 GHz J-band region. There also seems to be small ab-
sorptions at about 7 and 8 GHz. All three ferrite-treated resins show
similar effects, the 3/8"-thick composite with the impure 50-kg fer-
rite exhibiting the most pronounced. However, from 4 to 8 GHz a small
enhancement in attenuation of the 5-kg (pure) ferrite over the 50-kg
(impure) ferrite did seem to occur. In general, there was a negligible
dependence of absorption on the purity of the ferrite powder, suggest-
ing that the Fe content of the sludge could be significantly reduced
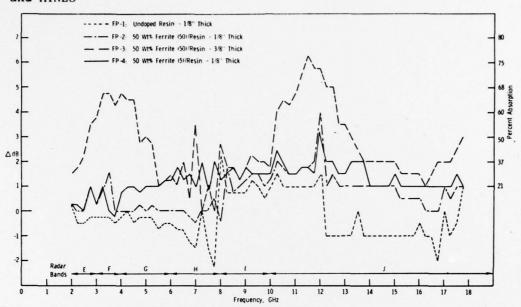without appreciably affecting the radar absorption properties.

*McCAULEY, HALPIN, EITELMAN
and HYNES

Figure 12. Graphical results of microwave absorption tests of ferrite/resin composites;
difference in amplitude from Al plate (ΔdB) plotted against frequency (GHz).

VIII. SUMMARY AND CONCLUSIONS

A comprehensive study has been carried out to evaluate the
microwave absorption potential of a ferrite sludge derived by a copre-
cipitation treatment of industrial waste water. Full characterization
of two different batches of sludge showed quite different chemistries,
especially in the amount of iron content; one was iron-rich, the other
iron-poor, differing by about 15% to 16% Fe. The iron-rich ferrite
was determined to be 100% ferrite spinel, whereas the iron-poor ferrite
was only 60% ferrite. The powders have an extremely fine grain size
with a surface area of 20 to 30 m^2/g, indicative of an average equiva-
lent spherical diameter of about 0.05 μm. Careful thermal analysis
and sintering studies clearly demonstrated that only carefully con-
trolled atmospheres can be used to sinter the ferrite sludge powder
into a mechanically coherent ferrite ceramic. The major problem is
the oxidation of the ferrite to the nonmagnetic α-Fe_2O_3 (hematite)
phase at relatively low temperatures. Because of these constraints a
ferrite/resin (EPON 828/Z) composite system was selected as the most
appropriate material configuration for the microwave absorption
evaluation of the sludge. Microwave testing of four test panels in
the Fort Huachuca anechoic chamber demonstrated significant attenuation
of the microwave signals by a 3/8"-thick ferrite/resin composite in two

regions: 3 to 4.5 GHz (68%) and 10 to 13 GHz (75%). There seemed to
be no significant difference in results using either the iron-rich or
iron-poor ferrite sludge material, showing that significantly less iron
could be used in the ferrite coprecipitation waste water treatment
process.

The data obtained in our investigation and the unconfirmed
results of NEC strongly suggest that the ferrite sludge concept could
prove to be an important new method of inexpensively reducing radar
signatures, by either a coating technique or smoke generation, while
stimulating the ecologically important purification of industrial
waste water. Also the possibility exists that the fundamental concept
confirmed in this program could be further optimized. Moreover, it is
our conclusion that the fabrication of massive fully sintered radar-
absorbing ceramic ferrites would prove prohibitively expensive and
unjustifiable from a practical point of view. Instead, the organic-
based matrix concept seems to be the most practical and feasible at
this point, as well as the use of the ferrite powder in camouflage
smokes.

IX. ACKNOWLEDGMENTS

*McCAULEY, HALPIN, EITELMAN
and HYNES

REFERENCES

1. ASAI, G. N. *Research Planned by Nippon Electric Co., Ltd., on the Prevention of Electromagnetic Wave Reflection by the Use of Ferrite Sludge.* U.S. Army Foreign Science and Technology Center, Report 2-253-0111-74, February 1974.

2. TAKADA, T., and KIYAMA, M. *Preparation of Ferrites by Wet Method: Ferrites* in Proceedings of the International Conference,.ed., Y. Hoshino, S. Iida, and M. Sugimoto, University Park Press, Tokyo, Japan, 1970.

3. PAULUS, M. *Preparation Conditions of the Ferrites* in Preparative Methods in Solid State Chemistry, ed., P. Hagenmuller, Academic Press, New York, 1972, p. 487-531.

4. GRAY, T. J. *Oxide Spinels* in High Temperature Oxides - Part IV, Refractory Glasses, Glass-Ceramics, and Ceramics, ed., A. M. Alper, Academic Press, New York, 1971, p. 77-107.

5. BLASSE, G. *Crystal Chemistry and Some Magnetic Properties of Mixed Metal Oxides with Spinel Structure.* Philips Research Reports Supplements, no. 3, 1964, 139 pages.

6. ASAI, G. N. *Research on Treatment and Properties of Ferrite Sludge for Microwave Absorption.* U.S. Army Foreign Science and Technology Center, Report 2-253-0593-75, 12 August 1975, (FOUO).

7. AKASHI, T., SUGANO, I., OKUDA, T., and TSUJI, T. *Sintering of Coprecipitated Manganese Zinc Ferrite Powder: Ferrites* in Proceedings of the International Conference, ed., Y. Hoshino, S. Iida, and M. Sugimoto, University Park Press, Tokyo, Japan, 1970.

8. AKASHI, T., SUGANO, I., KENMOKU, Y., SHINMA, Y., and TSUJI, T. *Low-Loss and High-Stability Mn-Zn Ferrites* in Proceedings of the International Conference, ed., Y. Hoshino, S. Iida, and M. Sugimoto, University Park Press, Tokyo, Japan, 1970.

THE PRODUCTION OF DECABORANE-14 FROM DIBORANE BY
LASER INDUCED CHEMISTRY (U)

*JAMES A. MERRITT, DR., HARRY C. MEYER, DR. RAYMOND I. GREENBERG, DR.,
and GEORGE A. TANTON, DR.
US ARMY MISSILE RESEARCH AND DEVELOPMENT COMMAND
REDSTONE ARSENAL, ALABAMA 35809

I. INTRODUCTION

High performance solid propellant fueled rocket motors re-
quire burning rate stabilizers to achieve fast burn rates. Presently,
n-hexylcarborane (NHC) is considered to be one of the most suitable
burning rate modifiers for solid propellant fuels. Its production in-
volves reacting 1-octyne with decaborane-14. The price and quantity
limiting factor in the supply of NHC is the lack of an industrial pro-
cess for synthesizing large quantities of decaborane inexpensively.

Thermal (pyrolysis) methods of influencing chemical process-
es lead, mainly, to the excitation of all degrees of freedom of the
molecule. Both external (translational) and internal (electronic,
vibrational, and rotational) degrees of freedom are usually in thermo-
dynamic equilibrium. In addition to there being an unproductive waste
of energy, reactions with equilibrium excited molecules characteris-
tically proceed in the direction of breaking the weakest bond, have a
considerable percent of back reaction, many side reactions, and pro-
duce polymers.

A new approach to the problem of chemical conversions of
substances would be to consider the possibility of influencing not a
molecule as a whole, but its individual bonds. Such a method of se-
lective excitation can be realized by means of lasers (1-5).

The recent development of high-power infrared lasers has
promoted a number of studies dealing with the interaction of this in-
tense radiation with organic and inorganic substrates. The often

1

selective character of the excitation, which can lead to isotope enrichment (6), and other unique features of this new type of photochemistry (7,8) have already been recognized and are attracting considerable interest. Earlier research in this laboratory using laser photochemistry to remove the trace impurity $COCl_2$ (phosgene) from boron trichloride, BCl_3 photosentized experiments, etc. (9) encouraged us to investigate the production of decaborane-14 from diborane.

Laser augmented chemistry of the following reaction leading to the production of solid decaborane-14 has been demonstrated:

$$10\ B_2H_6 + nh\nu \longrightarrow B_{10}H_{14} + 2\ B_5H_9 + 11\ H_2\ , \tag{1}$$

where h_ν can be either the CO_2 (00^01) - (10^00), R_{24}, $(978\ cm^{-1})$ or the DF $(1-0)$ P_{12}, $(2611\ cm^{-1})$ laser frequency.

II. EXPERIMENTAL

To avoid handling and storing the usual large (1 lb.) commercial quantities of B_2H_6 we used a 1% gaseous mixture of B_2H_6 in Ar supplied by the Linde Corporation. Immediately prior to irradiation, the B_2H_6 was recovered from Ar by passing the mixture into a condenser cooled to - 186ºC by an external bath of liquid Ar. A quantity of B_2H_6 sufficient for several irradiations was thus collected as a solid (F. P. = -126ºC) while the remaining Ar gas was pumped out of the system. Filling of cells and temporary storage of collected B_2H_6 was facilitated by transferring the B_2H_6 from the condenser to a cold storage bulb by allowing the condenser to warm up to ambient temperature. A cell was then filled by connecting it through a valve and quick disconnect fitting to the vacuum system and adjusting the storage bulb temperature to obtain the desired cell pressure. Typically brass or stainless steel cells 10 cm in length fitted with 3 cm diameter windows of NaCl, KCl, or KBr were used. The infrared spectrum (Figure 1) of the diborane did not reveal the presence of any impurities.

CO_2 laser irradiation was accomplished with both a CW grating tunable laser and a pulsed multigas tunable laser. The DF laser irradiation was accomplished with a 1 kW pulsed chemical laser by "Fresnelling off" 10-20 Watts. After irradiation for a given period of time and at varying powers the spectra of static products were obtained using both a Beckman IR5 spectrometer and a DIGILAB FTS-20B Fast Fourier Transform Spectrometer.
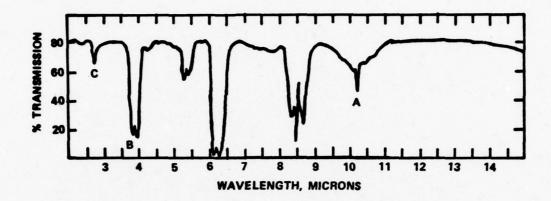
2

Figure 1. IR SPECTRUM OF GASEOUS B_2H_6

VIBRATIONAL BANDS RESONANT WITH   A: CO_2 LASER

                                         B: DF LASER

                                         C: HF LASER

III. RESULTS

<u>Low Power CO_2 Irradiation of B_2H_6 Resulting in the Formation of $B_{10}H_{14}$</u>

During irradiations with a CW CO_2 laser operating at wavelengths in the 975 cm^{-1} band of B_2H_6, Figure 2, we observed that white crystals formed along the gas cell walls when the power density of the laser did not exceed a constant threshold of 10 W/cm^2. Figure 3 is a photograph of crystals formed during a 30 minute irradiation of 400 Torr of B_2H_6 with the R24 laser line, which represents the optimum conditions found within the range of 10 to 800 Torr pressure of B_2H_6 at the attainable wavelengths indicated in Figure 2.
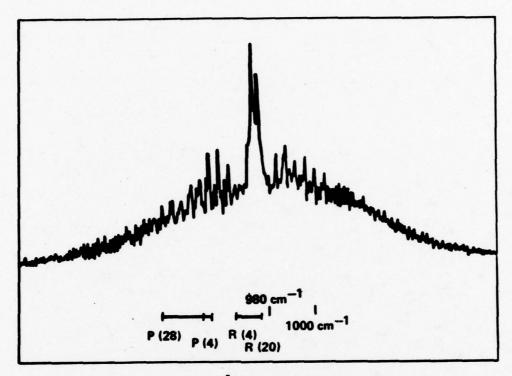
Figure 2. The ν_{14} (975 cm^{-1}) Fundamental Vibrational Band of B_2H_6 Showing the Region Irradiated with the (P_{28}-P_4) and (R_4-R_{20}) Lines of the CO_2 Laser.

In addition to white crystals the IR spectrum taken after irradiation, Figure 4 showed the presence of B_5H_9 with characteristic absorption peaks at 900, 1040, 1120, 1400, 1620, and 1800 cm^{-1}.

Prior to analysis, the crystals were removed from the cell by dissolving in MCH (methylcyclohexane) and later recrystalized in a controlled atmosphere dry box. Typically, 10 mg quantities were recovered after a 30 minute irradiation. A melting point determination revealed the crystals melt at 96°C. This together with the IR spectrum of the material in a NUJOL mull (Figure 5) showed the crystals to be pure $B_{10}H_{14}$.

Figure 3. $B_{10}H_{14}$ Crystals Formed when 400 Torr of B_2H_6 was Irradiated with the 975 cm^{-1} Frequency of a 10 W CO_2 Laser.

Characteristic $B_{10}H_{14}$ absorption peaks occur at 1950, 1900, 1520, 1020, and 720 cm^{-1} as shown in an absorption spectrum of commercial $B_{10}H_{14}$ in NUJOL mull, Figure 6. The characteristic BH stretching band is at 2600 cm^{-1}.

High Power CO_2 CW Irradiations of B_2H_6 Resulting in the Formation of Polymers

When the power density of incident laser radiation exceeded 10 W/cm^2, yellow or white colored particulates were formed immediately upon exposure of B_2H_6 to the radiation. Melting point determination showed the M.P. to exceed 300°C and the powder thus formed was insoluable in pentane and CCl_4.
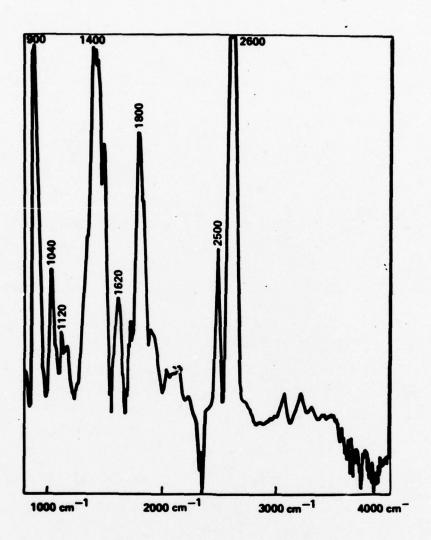
*MERRITT, MEYER, GREENBERG and TANTON

Figure 4.   IR Spectrum of the B_5H_9 Product
From CO_2 Laser Irradiated B_2H_6.

6

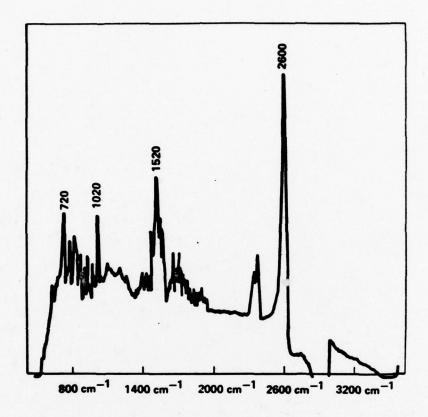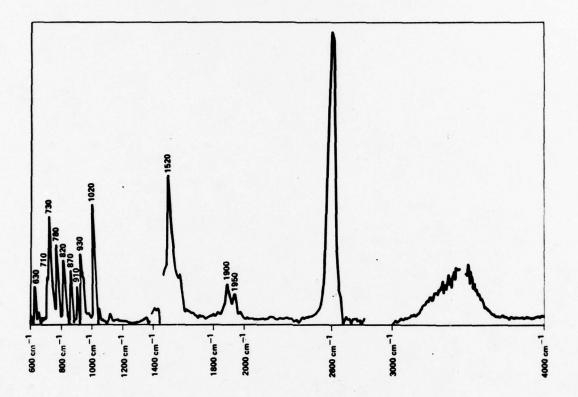Figure 5.   NUJOL Mull Spectrum of the Crystalline
White Solid Produced.

Figure 6.   NUJOL Mull Spectrum of $B_{10}H_{14}$.

8

*MERRITT, MEYER, GREENBERG and TANTON

Infrared spectra of the powder in KBr pellet gave broad band peaks at 3225, 2525, 1450, 1200 and 875 cm^{-1}. The mineral oil bands were observed at 3210, 2540, 1620, 1490, 1410, 1200, 830, 660, and 550 cm^{-1}, (Figure 7).

A further increase in intensity reaches a threshold where a yellow-white smoke appears at the entrance window and immediately inside it, and a yellow-white powder is deposited on the walls and throughout the cell. The intensity of the threshold for the production of this yellow-white powder is just above the intensity for the most efficient production of the decaborane crystals without the powder. It should be noted that the decaborane crystals are also produced when the powder is produced, but they are contaminated with the powder.

The thresholds for the production of decaborane and the powder were measured as a function of frequency. Measurements from the P_{20} and P_{40} to the R_{34} showed very little variation of these thresholds with frequency.
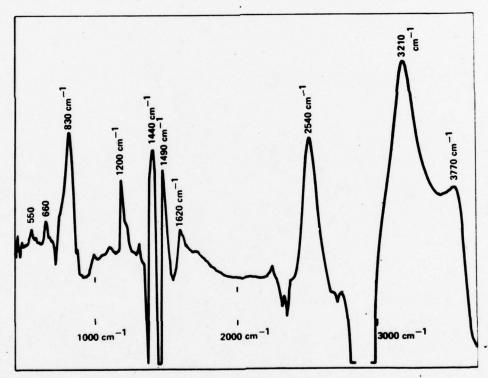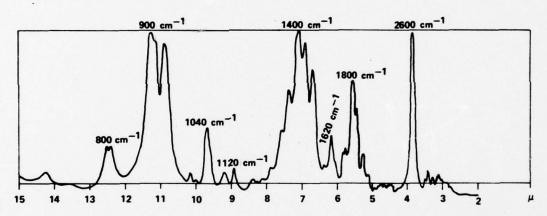

Figure 7. Spectrum of the Polymer Formed Dissolved in Mineral Oil.
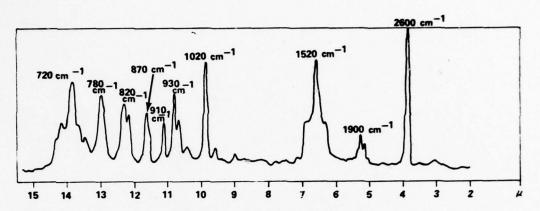
DF Laser Irradiations of B_2H_6 Resulting in the Formation of $B_{10}H_{14}$

When B_2H_6 (10 Torr in a 10 cm by 3 cm gas cell) was irradiated with the P_{12} (1-0 band) DF laser frequency (2614 cm^{-1}), clear crystals were deposited on the cell walls. The power diverted into the cell from the one kW laser beam was 13 Watts. The DF laser was a large pulsed (13 second duration) chemical laser and the target (cell) was placed some 100 yards from the laser. The laser beam entering the cell was a collimated beam one inch in diameter so that most of the volume of the cell was illuminated. The temperature rise of the cell wall was negligible (less than 5°C).

The spectrum of the cell contents revealed that the products observed were B_5H_9 (gas) and $B_{10}H_{14}$ (solid) (Figures 8 and 9).

LOW RESOLUTION IR SPECTRUM OF THE $B_5 H_9$ PRODUCT FROM DF LASER IRRADIATED B_2H_6

Figure 8

LOW RESOLUTION NUJOL MULL SPECTRUM OF $B_{10}H_{14}$ PRODUCT FROM DF LASER IRRADIATED B_2H_6

Figure 9

The production of decaborane-14 from diborane by laser in-
duced chemistry (LIC) (Equation 1) represents the first reported ex-
ample of the synthesis of a large polyatomic molecule by LIC. We
were unable to produce any decaborane-14 efficiently and of high puri-
ty using pulsed CO_2 laser excitation.

IV. DISCUSSION

From infrared and electron-diffraction evidence, diborane is
known to have the bridge structure (10, 11) belonging to the point
group D_{2h}, with a four membered B_2H_2 ring in a plane perpendicular to
the four terminal hydrogen atoms. The two fundamental vibrational
frequencies that are resonant with the CO_2 and DF lasers are shown
schematically along with their assignments and measured line positions
in Figure 10.

The infrared spectrum of diborane, Figure 1 , shows the vi-
brational bands that are resonant with the CO_2, DF and HF laser fre-
quencies. We have demonstrated the production of decaborane-14 using
both the CO_2 and DF laser frequencies. Diborane is mildly endotherm-
ic and the bridge bonds' force constant is about one-half as large as

11

○ HYDROGEN   ◯ BORON

ν_8

FUNDAMENTAL MODE OF

$B_2 H_3$

BH IN-PHASE STRETCHING

2614 Cm^{-1}

ν_{14}

FUNDAMENTAL MODE

OF $B_2 H_6$

BH$_2$ WAGGING

978 Cm^{-1}

THE LINE POSITIONS IN THE
PERPENDICULAR-TYPE BANDS OF DIBORANE

K"	2614 Cm^{-1} BAND P_Q	R_Q	K"	974 Cm^{-1} BAND P_Q	R_Q
0	–	2614.0	0	–	983.9
1	2608.5	2619.0	1	–	986.9
2	2604.5	2624.5	2	962.7	990.7
3	2599.5	2627.8	3	958.1	994.5
4	2594.5	2632.0	4	954.1	999.0
5	2590.5	2635.5	5	950.0	1003.4
6	2587.5	2639.5	6	946.0	
7	2581.5	2642.3	7	942.0	
8	2577.0	2647.5	8	937.9	
9	2573.0		9	933.0	
10	2567.0		10		
11					

SCHEMATIC OF THE FUNDAMENTAL VIBRATION MODES OF $B_2 H_6$
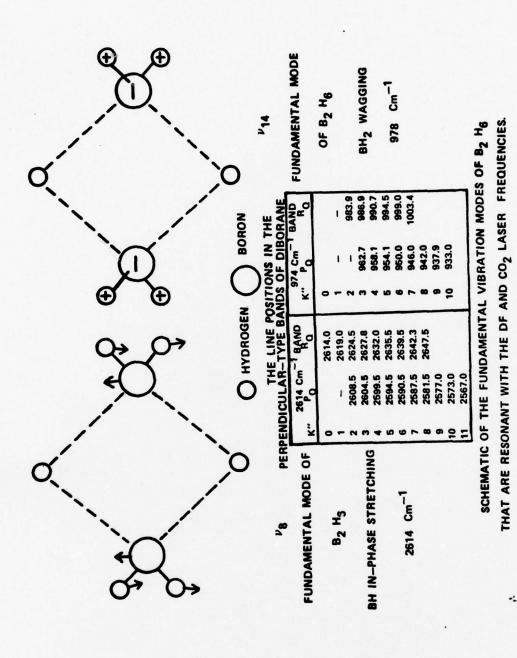THAT ARE RESONANT WITH THE DF AND CO$_2$ LASER FREQUENCIES.

Figure 10

12

that of the terminal hydrogen bonds. Yet there is substantial evidence that the initial reaction with diborane may be the breaking of the B-H (terminal hydrogen) bond.

The mercury-photosentized decomposition of diborane has been explained by postulating B_2H_5 as the primary intermediate, B-H ruptures being common in this type reaction (12). In isotope exchange studies, particularly deuterium, it is always the terminal hydrogen that is exchanged and not the bridged hydrogens. This is especially true for the higher boranes.

Many thermal studies have been conducted on diborane (13,14). The primary intermediate BH_3 has been postulated in these studies to account for B_5H_{11} and other boranes containing odd numbered boron atoms, and the dependency of the rate of product formation common to these reactions being of a one-half or three-halves power of the diborane concentration. However, no direct observance has yet been reported for BH_3. Another postulated mechanism is the molecular H_2 elimination from B_2H_6 as the primary photodissociation step giving $B_2H_4 + H_2$. Evidence for this consists of the large formation rate of H_2 (greater than the B_2H_5 formation rate) when low pressures of B_2H_6 are pyrolyzed in a hot-cold reactor (the inner cylindrical surface is 100°C while the outer surface is maintained at a much colder temperature), 90% of the product is observed to be B_4H_{10} ($B_2H_6 + B_2H_4$).

Considering the above the following reactions are reasonable:

$$B_2H_6 \xrightarrow{h\nu} 2\ BH_3 \tag{2}$$

$$B_2H_6 \xrightarrow{h\nu} B_2H_5 + H \tag{3}$$

$$B_2H_6 \xrightarrow{h\nu} B_2H_4 + H_2 \tag{4}$$

$$B_2H_4 + B_2H_6 \longrightarrow B_4H_{10} \tag{5}$$

$$2\ B_2H_5 \longrightarrow B_4H_{10} \tag{6}$$

$$2\ B_4H_{10} \longrightarrow B_8H_{12} + 4\ H_2 \tag{7}$$

$$B_8H_{12} + B_2H_6 \longrightarrow B_{10}H_{14} + 2\,H_2 \qquad\qquad (8)$$

$$B_4H_{10} + BH_3 \longrightarrow B_5H_9 + 2\,H_2 \qquad . \qquad (9)$$

V. CONCLUSIONS

Pure $B_{10}H_{14}$ can be produced in high yields at room temperature by laser induced chemistry (LIC). No polymer (which is always present in pyrolysis) was observed when we maintained the laser power below a certain threshold (10 W/cm^2 for the CO_2 laser). Also, the B_5H_{11} product is always present in pyrolysis; yet we never observed this product in LIC synthesis.

This research demonstrates that LIC is a new viable area of chemistry that can be used for cost-effective synthesis of large polyatomic molecules as well as smaller ones. Also, it demonstrates that LIC can be used in a prominent role for the synthesis of high energy binder molecules, burning rate modifiers and fuel additives with a substantial cost-savings (15).

14

REFERENCES

1. N. G. Basov, D. P. Markin, A. N. Oraevskii, and A. V. Pankratov, Dokl. Akad. Nauk. SSSR 198, 1043 (1971) (Sov. Phys. Dokl. 16, 445 (1971)).

2. N. G. Basov, D. P. Markin. A. N. Oraevskii, A. V. Pankratov, and A. N. Skachkov (to be published).

3. Moore, C. B., 1971, Ann. Rev. Phys. Chem. 22:387.

4. Moore, C. B., 1973, Advan. Chem. Phys. 23:41.

5. Moore, C. B., 1973, Acc. Chem. Res. 6:323; Moore, C. B., Zittel, P. F., 1973, Science 182:541

6. R. V. Ambartzumian and V. S. Letokhov, Acc. Chem. Res. 10, 61 (1977).

7. D. F. Dever and E. Grunwald, J. Am. Chem. Soc. 98, 5055 (1976).

8. W. Braun and W. Tsang, Chem. Phys. Lett. 44, 354 (1976).

9. J. A. Merritt and L. C. Robertson, J. Chem. Phys. 67, 3645 (1977).

10. W. C. Price, J. Chem. Phys. 16, 894 (1948).

11. K. Hedberg and V. Shomaker, J. Am. Chem. Soc. 73, 1482 (1951).

12. T. Hirata and H. E. Gunning, J. Chem. Phys. 27, 477 (1957).

13. R. P. Clark and R. N. Pease, J. Am. Chem. Soc. 73, 2132 (1951).

14. J. K. Bragg, L. V. McCarty, and F. J. Norton, J. Am. Chem. Soc. 73, 2131 (1951).

15. "The Laser Revolution in Energy Related Chemistry", Workshop, NSF, May 9-11, 1976.

15

METHOD TO EXPERIMENTALLY DETERMINE THE AERODYNAMIC PRESSURE DISTRIBUTION ON SPINNING BODIES

Miles C. Miller
Weapon Systems Concepts Team
US Army Armament Research and Development Command
Aberdeen Proving Ground, Maryland 21010

PREFACE

INTRODUCTION

Many Army weapon systems involve spinning aerodynamic bodies such as artillery shells, guided missiles, and Magnus rotor submunitions. The source of numerous flight instability problems has been related to the effect of body spin on the aerodynamic surface pressures. Although wind tunnel testing is the major means for determining the aerodynamic characteristics of these configurations, measurement of the surface pressures has been limited to non-spinning models. This paper describes a new and unique method to experimentally measure the aerodynamic pressure distribution on the surface of spinning wind tunnel models incorporating a special instrumentation scheme. A pressure tap located in the non-spinning inner portion of the wind tunnel model detects the surface pressure through a series of vent holes in the spinning outer portion of the model, the pressure being retained for measurement by means of a sliding seal arrangement. The feasibility of the method was thoroughly demonstrated by means of wind tunnel tests of fundamental model configurations.

The test program was conducted in two phases. First, the test method measured the surface pressure distribution acting on a spinning smooth surfaced cylinder in crossflow in the critical Reynolds number regime. This proved the method's capability to measure a steady state flow condition under various spin rate and free stream velocity conditions. The second phase involved surface pressure measurements on a spinning Magnus rotor. This extended the capability of the testing method to bodies

having irregular surface features and an unsteady, periodic flow field as compared to the smooth surfaced cylinder. Aerodynamic forces and flow field observations obtained by conventional testing techniques confirmed the validity of this unusual method for determining the pressure distribution on spinning bodies.

SYMBOLS

b span of model; $b = 8.482$ in

C_D drag coefficient $(D/q_\infty S)$

C_L lift coefficient $(L/q_\infty S)$

C_P pressure coefficient $(\Delta P/q_\infty)$

D drag force

d diameter of model; $d = 5.171$ in

L lift force

P pressure on surface of model

P_∞ free stream static pressure

q_∞ dynamic pressure $(\rho_\infty V_\infty^2/2)$

R_d Reynolds number $(V_\infty d/\nu)$

S reference area (bd); $S = .3046$ ft^2

t time ($t = 0$ when outer flat surface of upper driving vane is parallel to free stream velocity)

V_∞ free stream velocity

ΔP surface pressure referred to free stream static pressure $(P - P_\infty)$

α angle of attack

ρ_∞ air density

ν kinematic viscosity

$\emptyset$ circumferencial location on model at which pressure is being measured (angle between radial direction of pressure tap and free stream velocity)

Ψ model rotational attitude (angle between flat external surface of upper driving vane and free stream velocity)

Ω revolutions of model from $t = 0$ condition

ω spin rate

$\bar{\omega}$ tip speed ratio $(\omega d/2V_\infty)$

BACKGROUND

A non-spinning body travelling through air causes changes in the local air flow direction and velocity, producing an aerodynamic pressure distribution over the external surface of the body. This surface pressure results in net aerodynamic forces which influence the flight motion and trajectory of the body and represents the important middle stage between the flow field and the resultant aerodynamic forces as illustrated in Figure 1.

In similar manner, a spinning projectile also creates aerodynamic surface pressures which have led to the so-called Magnus effect. This aerodynamic phenomenon, due to the combination of body spin and attitude, produces forces and associated moments which have resulted in flight instability problems for several military projectiles (Ref 1). An understanding of the surface pressure distribution on spinning bodies could lead to identification of the source of the Magnus effect and body configurational modifications for its elimination.

Numerous unsuccessful attempts have been made to measure these Magnus associated pressures in wind tunnel tests utilizing a variety of techniques and instrumentation. The results described in this paper represent the first time that the aerodynamic surface pressure distribution on a spinning body has been obtained experimentally and validated by comparison with direct force measurements.

GENERAL APPROACH

The testing method evolved in this study is based on a unique model design and instrumentation arrangement. The model is composed of two parts: a non-spinning inner portion containing the pressure measuring instrumentation and a spinning outer portion representing the external aerodynamic surface of the body being evaluated. Figure 2 contains a schematic drawing of a cross-sectional view of the wind tunnel model looking along the longitudinal axis (i.e., spin axis) and illustrates the key elements of the test methodology. The stationary (i.e., non-spinning) model core contains a pressure tap oriented radially outward at an angle (∅) to the direction of the free stream velocity. The angle (∅) defines the circumferential location on the surface of the model at which the pressure is being measured. A thin walled, cylindrical shell is located concentrically around the core and is attached to the core by means of bearings located at each end. The shell is free to rotate or spin about the core and represents the external surface of the spinning model body. A small vent hole is located through the shell, such that it will line up with the face of the pressure tap once every revolution of the shell about the core. The gap between the face of the pressure tap and the inner surface of the shell is isolated (i.e., longitudinally and circumferentially) by means of a sliding seal attached to the outer end of the pressure tap. The cavity created within this seal will be open to the pressure acting

on the outside surface of the shell when the vent hole is aligned with the tap. When the vent hole rotates past the aligned position, the seal will cause the cavity to retain this pressure. Since the vent hole will be aligned with the cavity for a short time during each revolution, several revolutions of the spinning shell are required to have the pressure in the cavity reach a constant value equal to that acting on the surface of the shell. Tubing from the tap is routed through the model core to a pressure transducer and associated instrumentation located outside of the tunnel. Pressure measurements at various circumferential locations on the surface of the spinning body can be obtained by positioning the core and the attached tap at different attitudes ($\emptyset$) to the air flow.

MODEL DESCRIPTION

Details of the smooth cylinder wind tunnel model are shown in Figure 3. Circular end plates were located at each end of the cylinder to reduce tip flow effects. All of the pressure measuring elements in the model were contained within a radial hole located at the core midspan. References 2 and 3 describe the engineering design details associated with the model. The rubber "o" ring represents the most important element of the system in that it established the seal between the pressure tap cavity and the rotating cylindrical shell. Small coil springs pressed the "o" ring against the inner surface of the shell to affect a tight seal. An internal mechanism continuously deposited silicone grease onto the inner shell surface, for sealing and lubrication.

TEST ARRANGEMENT AND INSTRUMENTATION

The wind tunnel tests were conducted in the ARRADCOM Weapon Systems Concepts Team 28 x 40 inch open circuit, continuous flow subsonic wind tunnel. Figure 4 shows the model installed in the tunnel test section. A schematic of the instrumentation arrangement used in the wind tunnel tests is included in Figure 5. The model was mounted in the tunnel such that the longitudinal (i.e., spin axis) was in a vertical attitude and normal to the free stream velocity. Model spin was achieved by means of an electric motor mounted on top of the wind tunnel with a drive shaft extending to the top of the model. Model spin rate was indicated by a magnetic tachometer. The model core support strut was mounted to the tunnel turntable which could be rotated through 360 degrees, allowing the pressure tap to be set at any angle to the flow. The pressure in the cavity was transmitted by plastic tubing routed down the support strut to a transducer located outside the tunnel. The reference free stream static pressure was obtained by a static pressure probe in the tunnel test section. Both transducer and tachometer outputs were indicated as a function of time on a strip chart recorder.

TEST PROCEDURE AND DATA REDUCTION

The test procedure was to establish the desired air flow velocity in the tunnel test section and then use the spin motor to bring the model shell to the test spin rate value. Approximately 30 seconds were required for the pressure reading to equilibrate and be recorded. The core was then rotated to the next angular position while the model was spinning, this sequence being repeated at 10 degree increments around the entire circumference of the model surface. The pressure measurements were reduced to coefficient form and presented graphically. Data reduction techniques and associated equations are fully described in Reference 2. The pressure distribution was integrated over the surface of the model to obtain the resultant lift and drag coefficients as defined in Figure 6.

SMOOTH CYLINDER TEST RESULTS

Pressure distributions were obtained on the smooth cylinder model at various spin rates for three different tunnel velocities corresponding to subcritical, critical, and supercritical Reynolds numbers. The drag coefficient of the non-spinning cylinder in cross flow undergoes an abrupt change between a Reynolds number of 300,000 to 500,000, referred to as the critical Reynolds number regime. Figure 7 shows the drag coefficients as a function of free stream velocity for the non-spinning cylinder with end plates used in these tests. Note that the critical Reynolds number occurs at a tunnel velocity of 90 mph. Test velocities of 60 mph and 120 mph were selected as representative subcritical and supercritical test conditions, respectively.

Figure 8 contains a summary of the pressure distributions obtained during these tests. Note that at certain spin rates, the net lift force has an upward (positive) directional sense. For other spin and velocity conditions it is directed downward (negative). These pressure profiles provide a quantitative measure of the Magnus force and allow interpretation of boundary layer and flow separation effects. A detailed analysis of the pressure data obtained in this study can be found in Reference 2. The relation between the boundary layer, tip speed ratio, and Reynolds number on the resultant aerodynamic forces acting on a spinning cylinder under these flow conditions have been described qualitatively by other investigators (Ref 4). However, the data in Figure 8 is the first time that surface pressure distribution verification of these effects have been experimentally obtained.

A separate series of wind tunnel tests were conducted where the lift and drag forces acting on the model were measured directly by a force balance apparatus. These tests involved the same model and tunnel facility as the pressure measurement tests. Force coefficient values obtained by integrating the measured pressure profile data are shown in Figure 9. The close correlation between the conventionally obtained force data and

the integrated pressure measurements from this new testing method valid-
ates the pressure measurement method for this smooth model, steady flow
situation. Further, the primary benefit of this method is the detailed
insight into the source of the resultant forces provided by the pressure
distribution data.

MAGNUS ROTOR TESTS

A Magnus rotor is an asymmetrically shaped body that, when placed in
an air stream, will autorotate or spin, producing large aerodynamic lift
and drag forces. The Magnus rotor represents an aerodynamic configuration
with an irregular external surface as opposed to the smooth surfaced
cylinder previously tested. Also, unlike the steady flow field of the
smooth cylinder, the Magnus rotor generates an unsteady, periodic flow
field. The model used in the previous tests was modified to a cylindrical
Magnus rotor configuration by the simple addition of the four full span
driving vanes and additional vent holes. This configuration is represent-
ative of Magnus rotors and their aerodynamic characteristics.

Figure 10 contains a cross-section view looking along the longitudinal
axis of the Magnus rotor wind tunnel model. Because of the irregular ex-
ternal features, the surface pressure was measured at nine specific
locations relative to the external features. Due to the 90 degree rota-
tional symmetry, the model only required vent holes over a single quadrant.
Vent holes 8, 7, 6, 5, and 1 extended radially through the shell thickness
of .125 inches. Vent holes 5, 4, 3, and 2 had to pass through the thick-
ness of the driving vane resulting in a longer vent hole path, the longest
being vent hole 5 with a total length of .625 inch. This relatively long
length did not interfere with the ability to accurately acquire the sur-
face pressure.

The wind tunnel facility, model mounting, and instrumentation arrange-
ment were identical to that used with the smooth cylinder. The model was
spun by means of the spin motor to 2050 RPM providing the desired 0.46
steady state tip speed ratio for the tunnel velocity of 100.5 FPS. Only
one vent hole at a time was open for a specific test, the others being
closed off. The testing procedure was similar to that of the smooth
cylinder, with pressure data being obtained at 10 degree increments over
the entire circumference of the model surface.

Figure 11 contains selected examples of pressure data for three of the
vent hole locations. These data indicate the pressure acting at that par-
ticular location on the surface of the model over a complete 360 degree
rotation relative to the free stream direction. Similar plots were
obtained for all vent hole locations. At a specific rotational attitude
of the spinning model a particular vent hole location occurs at four
different circumferential positions relative to the free stream vector as
shown in Figure 12. Using data for all the vent holes, the complete sur-

face pressure distribution acting over the spinning Magnus rotor can be determined. Figure 13 shows the pressure distribution acting on the spinning Magnus rotor model over a quarter revolution at 11.25 degree increments. Note that for the 2050 RPM spin rate of the model, this corresponds to the pressure distribution at 0.9 milli-second intervals. This cycle is repeated four times during a complete 360 degree revolution. Once the basic data plots are obtained, pressure distributions can be determined over any time increment.

The resultant lift and drag coefficient were computed by integrating the pressure distribution over the model surface. These values are shown in Figure 13 and are presented as a function of model rotational attitude and time in Figure 14. Although the cyclic nature of the lift and drag of a Magnus rotor had been considered, this represents the first time that quantitative data have been obtained. Note that both the lift and drag are sinusoidal and out of phase. Because of the high rotational spin rates of Magnus rotors and the limited response of wind tunnel balance systems only the average aerodynamic forces can be measured in force type wind tunnel tests. The values for both the lift and drag derived from the pressure data show good correlation with the average values directly measured during separate force wind tunnel tests as also shown in Figure 14. The integrated pressure data, however, provide a more detailed insight into this cyclic aerodynamic effect. Complete details of the Magnus rotor tests are contained in Reference 5.

Flow visualization tests were conducted with the cylindrical Magnus rotor using a special smoke flow wind tunnel facility. High speed film records obtained from these tests were used to construct line drawings of the flow field streamlines for selected model rotational attitudes. Figure 15 shows these streamlines superimposed on the pressure distribution measured for the same conditions, providing a complete physical picture of the flow field, aerodynamic surface pressure distribution, and resultant aerodynamic forces for the spinning Magnus rotor at sequential instants of time. The unsteady, separated flow and in particular, the strong vortex shed off the retreating (upper) driving vane are graphically depicted. Note that this vortex produces a high velocity flow and corresponding low pressure over the forward portion of the vane. This figure illustrates the powerful tool provided by the pressure distribution toward interpreting and analyzing aerodynamic phenomenon on spinning configurations.

The flow over the Magnus rotor although unsteady, is periodic with model rotation. It is this latter factor that allows this pressure measuring method, which uses steady state instrumentation, to measure an unsteady flow condition. This same situation occurs in most non-transient flow fields around spinning bodies, thereby allowing this testing method to be used to investigate a variety of aerodynamic problems.

MILLER

FUTURE APPLICATIONS

The spinning body in cross flow represents an especially demanding condition for demonstrating the pressure measuring capabilities of this test method due to the severe influence of Reynolds number and model spin effects in the case of the smooth cylinder and the unsteady vortex flow for the Magnus rotor. The successful results indicate that the test method can be applied to a variety of model configurations and model orientations to the air flow (i.e., angles of attack). Because all of the instrumentation elements are internal to the model, the method could also be used in any Mach number regime including subsonic, transonic, and supersonic.

As the previous tests with the smooth cylinder will allow extension of the testing method to a variety of spinning models having smooth external shapes, so the Magnus rotor tests allow the method to be used with spinning models having irregular surface features as illustrated in Figure 16. Multiple pressure taps could be incorporated into a wind tunnel model at several locations along its length to provide the pressure distribution acting over the entire model surface. The most important features of this testing method are summarized as follows:

.All elements of instrumentation located within model or outside tunnel test section

.No inertial or dynamic loads acting on transducer

.Direct connection between pressure tap in model and instrumentation outside tunnel

.Constant pressure reading does not require high response transducer

.Rapid pressure surveys possible

.Applicable to any model orientation to free stream

.Can be used in any speed regime (subsonic, transonic, and supersonic)

.Will function with models having irregular external surface features

.Can be used with both steady flow or unsteady, periodic flow situations

CONCLUSIONS

1. This study has demonstrated the validity of a new testing method by which the aerodynamic pressure distribution acting on the surface of a spinning body can be measured in the wind tunnel.

2. Integration of the resulting pressure distribution data not only provides a quantitative measure of the Magnus forces but the pressure distribution provides a unique insight into the critical middle stage between the flow field and the resulting aerodynamic forces allowing a more accurate interpretation of boundary layer and flow separation effects.

MILLER

3. The test method can be utilized in both steady flow and unsteady periodic flow situations and is applicable to a variety of model external configurations, angles of attack, spin rates, and Mach number regimes.

REFERENCES

1. Jacobson, I. D., "AGARDograph No. 171. Magnus Characteristics of Arbitrary Rotating Bodies." November 1973.

2. Miller, M. C., "A Technique to Measure the Pressure Distribution Acting on the Surface of a Spinning Body in a Wind Tunnel," Edgewood Arsenal Technical Report ED-TR-76070. September 1976. Unclassified.

3. Miller, M. C., "Surface Pressure Measurements on a Spinning Wind Tunnel Model," AIAA Journal, Page 1669. December 1976.

4. Swanson, W. M., "The Magnus Effect: A Summary of Investigations to Date," Journal of Basic Engineering. Transactions of the ASME, Pages 461-470. September 1961. Unclassified.

5. Miller, M. C., "Wind Tunnel Measurements of the Surface Pressure Distribution on a Spinning Magnus Rotor," Proceedings AIAA 10th Aerodynamic Testing Conference. 19-21 April 1978.

Figure 1. AERODYNAMIC EFFECTS

a) FLOW FIELD

b) SURFACE PRESSURE

c) NET FORCES

Figure 2. TESTING METHODOLOGY

END VIEW

NOTE: DIMENSIONS IN INCHES

Figure 3. WIND TUNNEL MODEL DETAILS
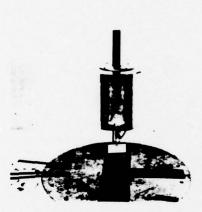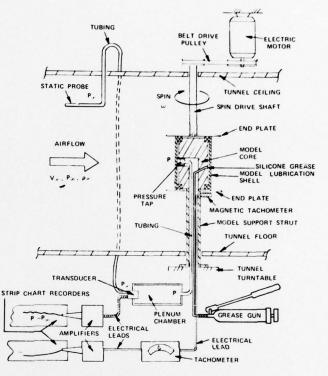
SIDE VIEW (A-A)

Figure 4. MODEL MOUNTED IN WIND TUNNEL
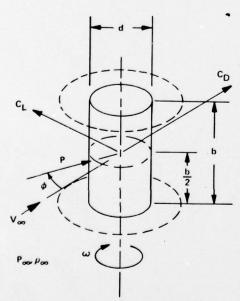

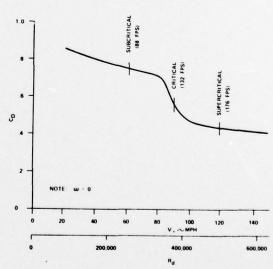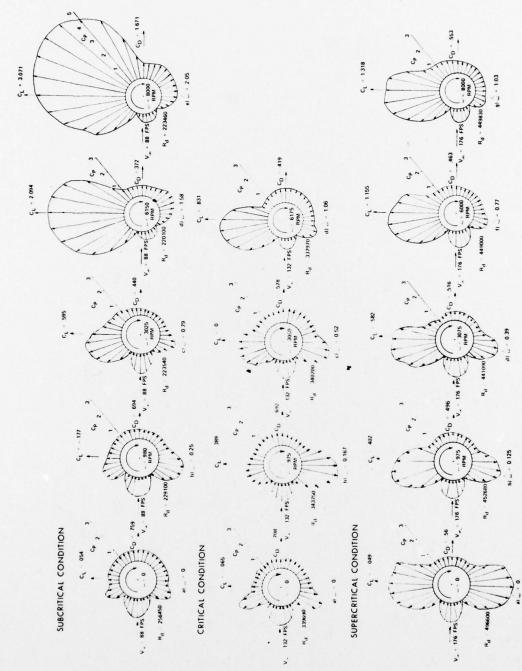Figure 5. INSTRUMENTATION ARRANGEMENT

Figure 6. DATA SYMBOLS

Figure 7. DRAG COEFFICIENT FOR NON-SPINNING SMOOTH CYLINDER MODEL IN CROSS FLOW

MILLER

Figure 8. SURFACE PRESSURE DISTRIBUTIONS ON SPINNING SMOOTH CYLINDER IN CROSS FLOW FOR VARIOUS FREE STREAM VELOCITIES AND SPIN RATES

VENT HOLE NO. 5

V_∞ = 100.5 FPS
$\hat{\omega}$ = 2050 RPM

VENT HOLE LOCATIONS

VENT HOLE NO. 4

VENT HOLE NO. 7

Figure 11. SURFACE PRESSURE DURING A COMPLETE MODEL REVOLUTION

a) PRESSURE AT VENT HOLE NO. 8 DURING A COMPLETE MODEL REVOLUTION
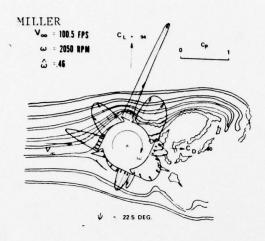
b) PRESSURES AT VENT HOLE NO. 8 AT A MODEL ROTATIONAL ATTITUDE OF ψ = 0°

c) PRESSURES AT VENT HOLE NO. 8 AT A MODEL ROTATIONAL ATTITUDE OF ψ = 45°

V_∞ = 100.5 FPS
$\hat{\omega}$ = 2050 RPM

Figure 12. PRESSURE DETERMINATION

V_∞ (FPS)	ω (RPM)	$\hat{\omega}$	R_d	DATA OBTAINED FROM PRESSURE TESTS		DATA OBTAINED FROM FORCE TESTS	
				C_L	C_D	C_L	C_D
88 →	0	0.	256450	.054	.759	0.	.755
	980	.25	229100	-.177	.694	-.180	.690
	3025	.79	223540	.595	.440	.560	.480
	6150	1.58	220100	2.094	.372	2.20	.620
	8000	2.05	223460	3.071	1.671	3.60	1.80
132 →	0	0.	339690	-.065	.708	0.	.700
	975	.167	343750	-.389	.692	-.530	.560
	3050	.520	340200	0.	.578	.06	.530
	6175	1.06	337970	.831	.419	1.11	.425
176 →	0	0.	496600	-.049	.560	0	.440
	975	.125	452680	.402	.496	.240	.480
	3075	.390	441090	.582	.516	.58	.510
	6000	.77	449000	1.155	.463	.73	.460
	8000	1.03	449830	1.318	.553	1.510	.600

NOTE: FORCE DATA ONLY OBTAINED UP TO 4500 RPM. COMPARISON OF DATA ABOVE THIS SPIN RATE IS BY EXTRAPOLATION.
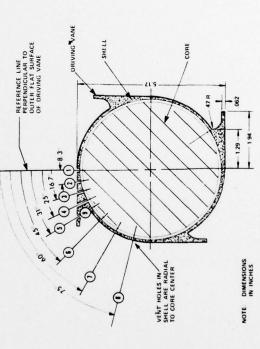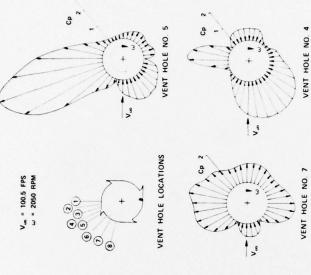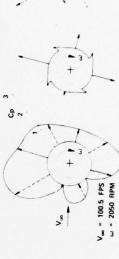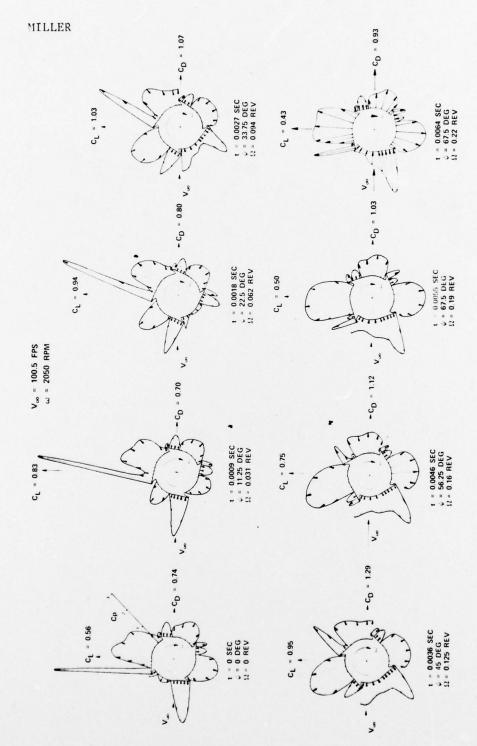
Figure 9. COMPARISON OF FORCE AND PRESSURE TEST DATA

REFERENCE LINE PERPENDICULAR TO OUTER FLAT SURFACE OF DRIVING VANE

DRIVING VANE

SHELL

CORE

VENT HOLES IN SHELL ARE RADIAL TO CORE CENTER

NOTE: DIMENSIONS IN INCHES

Figure 10. MAGNUS ROTOR MODEL CONFIGURATION
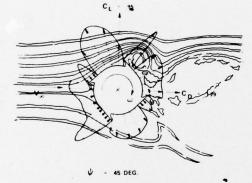
Figure 13. SURFACE PRESSURE DISTRIBUTION ON SPINNING MAGNUS ROTOR IN CROSS FLOW AT SPECIFIC TIME INTERVALS

MILLER

V_∞ = 100.5 FPS
ω = 2050 RPM
$\hat{\omega}$ = .46

C_L = 94

0   C_P 1

C_D

ψ = 22.5 DEG.

C_L = 95

C_D

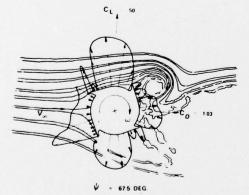ψ = 45 DEG.

C_L = 50

C_D = 1.03

ψ = 67.5 DEG.

Figure 15. FLOW FIELD, SURFACE PRESSURE DISTRIBUTION, AND RESULTANT AERODYNAMIC FORCES ON SPINNING MAGNUS ROTOR AT VARIOUS ROTATIONAL ATTITUDES
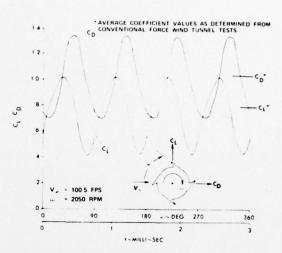
*AVERAGE COEFFICIENT VALUES AS DETERMINED FROM CONVENTIONAL FORCE WIND TUNNEL TESTS

C_D

C_D^*

C_L^*

C_L

V_∞ = 100.5 FPS
ω = 2050 RPM

C_L

V_∞ C_D

0   90   180 $\sim$DEG   270   360

0   1   2   3

t $\sim$ MILLI $\sim$ SEC

Figure 14. LIFT AND DRAG COEFFICIENT AS A FUNCTION OF TIME FOR SPINNING MAGNUS ROTOR

ω

V_∞

SMOOTH CYLINDER

ω

V_∞

MAGNUS ROTOR

α ω

V_∞

PROJECTILE WITH CONVENTIONAL BOATTAIL

α ω

V_∞

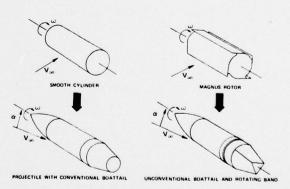UNCONVENTIONAL BOATTAIL AND ROTATING BAND

Figure 16. FUTURE APPLICATIONS FOR SURFACE PRESSURE TESTING METHOD